

FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGAGGACCTGGAGCTCCGGCTGCGTCTTCCCG
CAGCGCTACCCGCCATGCGCTGCGCGCCGGGCCGCGCTGGGGCTCTGCGCGTTCTGCTG
CTGCTGCGCCCGCGCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGT
GGTGGACAAGTTTAACAGGGGATGGTGGACACCGCAAAGAAGAAGTTTGGCGGCAGGAAACA
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATC
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCTGTGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC
CAGGGCGGATCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGGCGACGGGTCTGCGCGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAACCTGTGTGAGGAAAAACG
AAAAGTGTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCTGTGACGGCTTCGAAGAAACG
GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTTATTAGAAGGATGTCC
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCTGCACTGACAGCGCGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTGTGCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA
AACAGACTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTACAATAAAGCAATAGCA
TCACAAATTTACAAATAAAGCATTTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAAACT
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATTGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAATTAGGGTGTGGAAGTCCCCAGGCTCCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTT

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLPPAPEAAKPTPCHRCRGLVDKFNQGMVDATAKNFNGGNTANEKTLISKYESSSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWMLQLKSEYPDLFEWFCVKTILKCCSPGTYPGPDCLACQGGGSRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESKCTCSGLINRDCGECEVGWVLDE
GACVDVDECAAEPPPCSAQAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECTISGYAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSIVCVCPDGFEEDEDACVPPAEAEATEGESPTQLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343
and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and
313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

FIGURE 3

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCTCCCGGGGATCCTCTAGAGATCCCTC
 GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCCAGCCGTCTAAACGGGAACA
 GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG
 GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCTGAGCAGC**ATG**CCCCGAGGAGCGCCTTC
 CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCTCTGTGGCACTCGGGCGGAGGC
 CGGCCCGCCGAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA
 TAGGATTTTGAAGAAGATATCTTGATTGTTTCAGAGGGGAAAATGGACCTTTTACACATGAT
 TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC
 CTGGCAAGCTGCAGGGCAGGCAGAAATACCTTCTATGAATTCCTGTCTTGGCTCCTCGGATA
 AAGGCATCATGGCAGATCCAAACCGTCAATGTCCCTCTGCTGGGAAACAGTGCCTCACAAGGCA
 TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAACACAGGATGGGGTGGCAGCATTGAAGT
 GGATGTGATTGTTATGAATTTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT
 TCTTTAAACATGTCAACAAGCTGAGTGCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAAT
 GAAAGACGCATCTGCGAGTGTCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTG
 TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG
 GATTCTATGGAGTGAACCTGTGACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACC
 TGTTTCTACCTTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG
 CAAATGCCACAACCTCTGTGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAAGTGT
 CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT
 GGAACCTGCCATGAACCCAACAATGCCAATGTCAAGAAGGTTGGCATGGAAGACATGCACAA
 TAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCCGCCAGCTCAGGCAGC
 ACACGCCCTTCACTTAAAAAGGCCGAGGAGCGCGGGATCCACCTGAATCCAATTACATCTGG
TGAACTCCGACATCTGAAACGTTTAAAGTTACACCAAGTTCATAGCCTTTGTAAACCTTTCA
 TGTGTTGAATGTTCAAATAATGTTTACCTTAAAGAATACTGGCCTGAATTTTATTAGCT
 TCATTATAAATCACTGAGCTGATATTTACTCTTCTTTAAAGTTTCTAAGTACGCTCTGTAG
 CATGATGGTATAGATTTTCTGTTTCAGTGTCTTGGGACAGATTTTATATTATGTCAATTGA
 TCAGGTAAAAATTTTCACTGTGTAGTTGGCAGATATTTTCAAATTACAATGCATTATGGT
 GTCTGGGGGCGGGGAACATCAGAAAAGGTTAAATTGGGCAAAAATGCGTAAATCAAGAAAT
 TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTTCAGATTTTATTGTGAGATTTTAGAT
 GTTTGTTACATTTTAAAAAATTGCTCTTAAATTTTAACTCTCAATACAATATATTTTGGACC
 TTACCATTATTCAGAGATTTCAGTATTAACAAAAAATAATACACTGTGGTAGTGGCATT
 AAACAAATATAATATATTTCAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT
 TGGCTTGAAGCAATATAATATATTTGTAACAAAAACACAGCTCTTACCTAATAACATTTTAT
 ACTGTTTGTATGTATAAATAAAGGTGCTGCTTTTAGTTTTTGGAAAAAATAAATAAATAA
 AAAAAAATAAATAAATAAAGGCGCGCCGCGACTTAGAGTCAGCTGCAGAAGCTTGGC
 GCCTATGGCCCACTTGTATTATGCAGCTTATAATG

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 4

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2
MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSPVC
EPGCGAHGTCHEPNKCQCQEGWHRHCNKR YEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW
```

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTGCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG
 CCCAGCCACACCTTCCACAGGGCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC
 TGTGTTTGTCTGCCGTGGCTGGCCACTTGGCTCTGGGTGCCAGCAGGGTCGTGGGCGCCGG
 GAGCTAGCACCAGGGTCTGCACCTGCGGGGCATCGGGACGCGGGAGGCCGGTACTGCCAGGA
 GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT
 GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
 TGCCTCGGCGTGCCACCCCTTTTCCCGCGATCCAAGGATGTATGCATGGAGGTCTGTATCTTA
 TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT
 GGCATGTGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGGCTGG
 GAACCACAGCGCCTTCTGGGGCATGACCTGGATGGAGGGCAATTGCTACCGCCTGGGCACCA
 TCCGCCCATCTTCTCGGTATGAAACATGCATGAAATTTATACAGTGTGTAACCCAGGGGAG
 GTGCTTCCACAGCCTTCGAGGGCCTCTGAGAAGTGGCCCAACCTGATTTCATGAGCCTCTTGA
 CCAAGGCAACTGTGTGAGGCTCTTGGGCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
 CAATCCATTCTCTGGGACACATGACGCTGTCTGTGCGCCCAAGAACCTGCTGTCTTGTGAC
 ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGTCTGTGTCTGCGCTCG
 CCGAGGGGTGGTGTCTGACCACTGTCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC
 CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTGGGGCAAGCGCCAGGCCACTGCC
 CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACAGGTCACTCTCTGTCTACCGCCT
 CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA
 TGGAGGTGCATGAGGACTTCTTCTCTATACAAGGAGGCATCTACAGCCACACGCCAGTGTGAGC
 CTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG
 AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGAATGCGGCCAATCTTGGGGCC
 CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCAGCGCGTCAATGAGTGCACATCGAG
 AGCTTCGTGCTGGGCGTCTGGGGCGCGTGGGCATGGAGACATGGGTCTACTGAGGCTG
 CGGGCACCACCGGGGTCCGGCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCAATG
 GGGCGGTGACCCAGCCTCGCCCGACAGAGCCCGGGCGCAGGCGGGCGCCAGGGCGCTAAT
 CCGGGCGCGGTTCCGCTGACGACGCGCCCGCCTGGGAGCCGCGGCAGGCGAGACTGGCG
 GAGCCCCCAGACCTCCAGTGGGGACGGGGCAGGGCTTGGCCTGGGAAGAGCACAGCTGCAG
 ATCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC
 CCCAATACCCACCCCAATCCCGTATTCTTTTCTTTTATAGACAGGGTCTTGTCTCCG
 TTGCCAGGTTGGAGTGAGTGGCCCATCAGGGCTCACTGTAACTCCGACTCTTGGGTCA
 AGTGACCTCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCAACCACTGGC
 TAAATTTTGTATTTTTGTAAAGAGGGGGTCTCACTGTGTTGCCAGGCTGGTTTCGAAC
 CTTGGGCTCAAGCGGTCCACCTGCCTCCGCTCCCAAAGTGTGGGATTGCAGGCATGAGCC
 ACTGCACCCAGCCTGTATTCTTATCTTTCAGATATTATTCTTTTCTTTCACTGTTTAAAA
 TAAACCAAAGTATTGATAAAAAAAA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 6

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18359, pI: 7.45, NX(S/T): 1
MWRCPGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCEQDLCCRGRADDC
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPPPIQGCMHGGRIYPVLGTYWDNCNR
CTCQENRQWHGGSRHDSHQPGQLWLAWGEPQRLLGHDPG

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

[illegible]

AGGCTCCTTGGCCCTTTTTCACAGCAAGCTTNTGCNATCCCGATTCTGTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCCGCGNTGGGCGTCTCGATGGTGCTGTGTGTTCTCTGCGTGCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGTTCGGGGCAAGCGCCAGGCCACTGCCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCTCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATTAAGGAGCTGGTGGAGAATTGGCCCTGTCCAAGCCCTCATGGAG
TGCATGAGGAGACTTCTTCCCTATAAAGGGAGGCGACATGACGCACACGCCAGTGAAGCTTGGG
AGGCCAGAGAGATACCGCCGCGATGGGACCCATCAG

Year	Age	Sex	Location	Length (mm)	Weight (g)	Stomach contents	Notes
1961	10	M	St. Lawrence	100	10	Small crustaceans	
1962	12	F	St. Lawrence	120	15	Small crustaceans	
1963	14	M	St. Lawrence	140	20	Small crustaceans	
1964	16	F	St. Lawrence	160	25	Small crustaceans	
1965	18	M	St. Lawrence	180	30	Small crustaceans	
1966	20	F	St. Lawrence	200	35	Small crustaceans	
1967	22	M	St. Lawrence	220	40	Small crustaceans	
1968	24	F	St. Lawrence	240	45	Small crustaceans	
1969	26	M	St. Lawrence	260	50	Small crustaceans	
1970	28	F	St. Lawrence	280	55	Small crustaceans	
1971	30	M	St. Lawrence	300	60	Small crustaceans	
1972	32	F	St. Lawrence	320	65	Small crustaceans	
1973	34	M	St. Lawrence	340	70	Small crustaceans	
1974	36	F	St. Lawrence	360	75	Small crustaceans	
1975	38	M	St. Lawrence	380	80	Small crustaceans	
1976	40	F	St. Lawrence	400	85	Small crustaceans	
1977	42	M	St. Lawrence	420	90	Small crustaceans	
1978	44	F	St. Lawrence	440	95	Small crustaceans	
1979	46	M	St. Lawrence	460	100	Small crustaceans	
1980	48	F	St. Lawrence	480	105	Small crustaceans	
1981	50	M	St. Lawrence	500	110	Small crustaceans	
1982	52	F	St. Lawrence	520	115	Small crustaceans	
1983	54	M	St. Lawrence	540	120	Small crustaceans	
1984	56	F	St. Lawrence	560	125	Small crustaceans	
1985	58	M	St. Lawrence	580	130	Small crustaceans	
1986	60	F	St. Lawrence	600	135	Small crustaceans	
1987	62	M	St. Lawrence	620	140	Small crustaceans	
1988	64	F	St. Lawrence	640	145	Small crustaceans	
1989	66	M	St. Lawrence	660	150	Small crustaceans	
1990	68	F	St. Lawrence	680	155	Small crustaceans	
1991	70	M	St. Lawrence	700	160	Small crustaceans	
1992	72	F	St. Lawrence	720	165	Small crustaceans	
1993	74	M	St. Lawrence	740	170	Small crustaceans	
1994	76	F	St. Lawrence	760	175	Small crustaceans	
1995	78	M	St. Lawrence	780	180	Small crustaceans	
1996	80	F	St. Lawrence	800	185	Small crustaceans	
1997	82	M	St. Lawrence	820	190	Small crustaceans	
1998	84	F	St. Lawrence	840	195	Small crustaceans	
1999	86	M	St. Lawrence	860	200	Small crustaceans	
2000	88	F	St. Lawrence	880	205	Small crustaceans	
2001	90	M	St. Lawrence	900	210	Small crustaceans	
2002	92	F	St. Lawrence	920	215	Small crustaceans	
2003	94	M	St. Lawrence	940	220	Small crustaceans	
2004	96	F	St. Lawrence	960	225	Small crustaceans	
2005	98	M	St. Lawrence	980	230	Small crustaceans	
2006	100	F	St. Lawrence	1000	235	Small crustaceans	
2007	102	M	St. Lawrence	1020	240	Small crustaceans	
2008	104	F	St. Lawrence	1040	245	Small crustaceans	
2009	106	M	St. Lawrence	1060	250	Small crustaceans	
2010	108	F	St. Lawrence	1080	255	Small crustaceans	
2011	110	M	St. Lawrence	1100	260	Small crustaceans	
2012	112	F	St. Lawrence	1120	265	Small crustaceans	
2013	114	M	St. Lawrence	1140	270	Small crustaceans	
2014	116	F	St. Lawrence	1160	275	Small crustaceans	
2015	118	M	St. Lawrence	1180	280	Small crustaceans	</

GCTGCTTGCCCTGTGTGATGGCAGGCTTGGCCCTGCAGCCAGGCATGCCCCGTCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCGCGCAGTTGGCCCTCTGACCCTCATCAGCAAAGGCTG
CAGCTTGAAGTGCCTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCAGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACTGCGGTGTGGTGCCCAAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCTGTGTTCTTGAGGCACATCCTAAACGCAAGTCTGACCATTGTATGT
CTGCACCCCTGTCCCCACCCCTGACCCTCCCATGGCCCTCTCCAGGACTCCACCCCGGAGA
TCAGTCTTAGTGACACAGATCCGCCCTGCAGATGGCCCCCTCAACCCCTCTCTGCTGCTGTTT
CATGGCCCAGCATTCTCCACCCTTAACCTGTGCTCAGGCACCTCTTCCCCAGGAAGCCTT
CCCTGCCCAACCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGACTCAGGAGGGGCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTGCAGCTGAGTTCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACAATTCTGTGGGGCTCCTGTAATGGCAGCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAA

FIGURE 9

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLLPPSLRCSLHSACCSGDPASYRLWGAPLQPT
LGVVVPQASVPLLLTDLAQWEPVLVPEAHPNASLTMVYCTPVPHPDPPMALSRTPTRQISSSDT
DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSFAPSRGQALRRAQ

Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

FIGURE 10

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCGCTGCTGCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAAGTACACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCAGCGGGAAGAGCAAAGACTG
CGTGTTCACGGAGATCGTGTCTGGAGAACAACATATACGGCCTTCCAGAACGCCCGGCACGAGG
GCTGGTTTCATGGCCTTCACGCGGCAGGGGGCGGCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCCACGC
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCACCCGCCGACCAAGCGCACAC
GGCGGCCCCAGCCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCCCTCCCT
CACCCCTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCTTGAGGGCCGCGAAGCATCCGAGCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGACGGGTGGCAGGCCCTGGAGAGGAATG
GAGTGTACCCCTGATCTCAGGCCACCAGCCTCTGCCGGCTCCAGCCGGGCTCCTGAAGCC
CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAAGACAACCGTCTGGAGGTGGCTGTCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCAGCCCCAAACTCCTCTGGCTAGACTGTA
GGAAGGGACTTTTGTTTGTTGTTTTCAGGAAAAAAGAAAGGGAGAGAGAGAAAAATAG
AGGGTTGTCCACTCCTCACATTCACGACCAGGCCTGCACCCACCCCCAACTCCAGCCC
CGGAATAAAACCATTTTCCTGC

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 11

MGAARLLPNLTLCLQLLI LCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNRGKLGKPSGKSKDCVFTEIVLENNYTAFQ¹NARHEGWMAFT²RQGRPRQASRSRQ³NQREAHFIKRLYQGQL⁴FPFNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCCTCCTGTTGAATTTTTTGCACATGGAG
 GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
 CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTCTGGACTTCAACAGAACCCCATCCGAT
 CATTTTGATTTTGTGTTTATTTTTTTTTTCTTTTCTTTTCCCACCACATTGTATTTTAT
 TTCCGTACTTCAGAAATGGGCGCTACAGACCACAAAGTGGCCCGAGCCATGGGGCTTTTTCCT
 GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCAGAGGTGTCCAAACTCCTGGCCGTGCC
 CTAGTGTGTGCCGTGCGACAGGAACTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
 CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAATTAATAATGC
 TGGATTTCCTGCAGAACTGCACAAATGTACAGTCTGGTGCACACGGTCTACCTGTATGGCAACC
 AACTGGACGAATTCCTCATGAACCTTCCCAAGAATGTGAGAGTCTCCATTGTGAGGAAAAC
 AATATTTCAGACCATTTTCACGGGCTGCTCTTGCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
 GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCTTCCGGGAGGCTATTAGCC
 TCAAATTTGTTGTTTTGTCTAAGAAATCACCTGAGCAGTGTGCTCTTGGGCTTCTCTGTGGAC
 TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGTCTGCATATCCGACATGGCCCTCCAGAA
 TCTCAGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCTGACCAACAAGGGTATCGCCG
 AGGGCACTTACGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
 CACCCTCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTCGAGGACAAACAGAT
 AAACACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
 ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
 ACTGCTCGGAATAACCTTGGTTTTGTGACTGCAGTATTAATGGGTACAGAAATGGCTCAA
 ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCTGAACAAGTCCGGG
 GGATGGCCGTGAGGGAATTAATATGAATCTTTTGCTCTGCCACACGACCCCGCGCTG
 CCTCTCTTCAACCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCACCCCTCTCTAT
 TCCAAACCTGACAGAAGCTACACGCCTCCAACCTCTACACATCGAACTTCCCACGATTCT
 CTGACTGGGATGGCAGAGAAAGAGTGACCCCACTATTCTGAAGGATCCAGCTCTCTATC
 CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCAACCGTGATGGCATA
 CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTACAGGAGCGCATAG
 TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCGATCCACCTATCGGATT
 TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACCACTTTGTTACAGAGGC
 CACCACCATGCCCTCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
 CGTCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
 GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGTCATATGCACAAAAGGGGCGCTACACCTC
 CCAGAAGTGGAAATACAACCGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
 AGGACAACCTCCATCCTGGAGATGACAGAAACAGTTTTCAGATCGTCTCTTAAATAACGAT
 CAACCTCTTAAAGGAGATTTTCAGACTGCAGCCATTACACCCCAATGGGGCATTAATTA
 CACAGACTGCCATATCCCCAACACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
 ACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
 CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
 TTGTGCATTGGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG
 CTATCTTTCTATTTCAGTTAATTACAACAGTTTTGTAACTCTTGTCTTTTAAATCTT

100
 200
 300
 400
 500
 600
 700
 800
 900
 1000
 1100
 1200
 1300
 1400
 1500
 1600
 1700
 1800
 1900
 2000
 2100
 2200
 2300
 2400
 2500
 2600
 2700
 2800
 2900
 3000
 3100
 3200
 3300
 3400
 3500
 3600
 3700
 3800
 3900
 4000
 4100
 4200
 4300
 4400
 4500
 4600
 4700
 4800
 4900
 5000
 5100
 5200
 5300
 5400
 5500
 5600
 5700
 5800
 5900
 6000
 6100
 6200
 6300
 6400
 6500
 6600
 6700
 6800
 6900
 7000
 7100
 7200
 7300
 7400
 7500
 7600
 7700
 7800
 7900
 8000
 8100
 8200
 8300
 8400
 8500
 8600
 8700
 8800
 8900
 9000
 9100
 9200
 9300
 9400
 9500
 9600
 9700
 9800
 9900
 10000

FIGURE 13

MGLQTTKWP SHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFP MNLPKNVRVLHLQENNIQTI
SRAALQALLKLELHLLDDNSISTVGVEDGAFREAI SLKLLFSLSKNHLSSVPVGLPVDLQELR
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
PWFCDCSIKWVTEWLKIPSSLNVRGFMCGQGEQVRGMARELNMMNLLSCPTTTPGLPLFTP
APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWGDERVTPPISERIQLSIHFVND
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQBRIVSGEKQHLSLVNLEPRSTYRICLVPL
DAFNRYRAVEDTICSEATTHASYLNNGSNTASSHEQTTS HSMGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNNDQLLKG
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 14

ACTTGGAGCAAGCGGGCGGCGGAGACAGAGGCAGAGGCAGAAAGCTGGGGCTCCGCTCTCGCCTCCACGAGCG
ATCCCCGAGGAGAGCCCGCGCCCTTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTCGCCCTTCGC
TCGCTTCCAGGCGCGCGGGCTGCAGCCTTGGCCCTTGTCTCGCCTTGA AAAATGGGAAAAGATGCTCGCAGGCT
GCTTTCGTGTATCTCGGACAGATCGTCTCTTCCTGCCGAGGCGCAGGAGCGGTTCAGTGGGAGGTCCATCT
CTAGGGGCGAGACAGCTCGGACCCGACCGCAGACGGCCCTTCGGAGAGTTCCTGTGAGAACAGCGCGGCGAGACC
TGGTTTTTCATCATGTGACAGCTTCGCGAGTGTCAACACCCATGACTATGCAAAAGGTCAAGAGTTTATCTGTGGACA
TCTTGC AATTTCTTGACATTTGGTCTGTATGTCAACCGAGTGGGCTGCTC AATATGGCGCACTGTCAAGAATG
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTCTCAAGAGGATGCGGCATCTGTCTCACGG
GCACCATGATGCTGGGCTGGCCATCCAGTATGCCCTGAAACATCGCATTCAGAGAGCAGAGGGGGCCCGGCCCTTGA
GGGAGAATGTGCCACGGGTATCAATGATCGTGACAGATGGGAGACCTCAGGACCTCCGTGGCGGAGGTGGCTGCTA
AGGCACGGGACACGGGCTCTTAATCTTGGCATTTGGTGTGGGCGAGGTAGACTTCAACACTTGAAGTCTCAITG
GGAGTGAGCCCCATGAGGACCATGTCTTCTTGTGGCCAAATTCAGCGCAGATTGAGACGCTGACCTCCGTGTTCC
AGAAGAAGTTGTGCAACGGCCCATGCTGCAGCACCTTCGAGCATAACTGTGGCCACTTCTGATCAACATCCCTG
GCTCATACGCTCTGCAAGGTGCCAAACAGGCTACATCTCAACTCGGATCAGACGACTTGCAGAAATCCAGGATCTGT
GTGCCATGGAGGACCACAAGTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGTCTACAGTG
CTACGCCCTTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGACTACTGTGCCTCAGAAAAACCGGATGTGAAC
GTGAGTGTGTAAATGCTGATGGCTCTTACCTTTGGCAGTGGCATGAAGGATTTGTCTTTAAACCCAGATGAAAAA
CGTGCAACAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGGCTCAACATGGAGGAGAGCT
ACTACTGCCGTGCCACCTGGGCTCACTCTGGAACCCAAATGGCAAAACCTGCAGCCGATGGGACACTGTGCA
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGTCTCAGAAGSCTTCC
TCATCAACGAGGACCTCAAGACCTGCTCCCGGTGGATTAAGTCTGCTGTGAGTGAACATGCTTTGTGAATATCTCT
GTGTCAACATGAGCAGATCTTTTGGCTGTCACTGTGAGTGTCTTGGAGGACACGTGCTCCGCGAGCGATGGGAAGCGTGTG
CAGTAATTTGGAACCTTTGTGCTCTGGGGGACACGGTGTGAAACATTCGTGTGTAGCAGTGAAGATTCGTTGTGT
GCCAGTGTCTTGAAGGTTTATATATCTCCGTGAAGATGGAAAAACCTGCAGAAGGAAGATGTCTGCCACGCTATAG
ACATGAGTGTGACACACATTTTGTGTGAACAGTGAAGCATCAACACGTCGCGAGTGTCTGGAGGGATTCGGGCTCG
CTGAGGATGGGAAACGCTGCCAGGAAGGATGTCTGCAAAATCAACCCACCATGGCTGCGAGACATTTTGTGTGA
ATAATGGGAATCTCATCTGCAAAATGCTCAGAGGGATTTGTCTTAGCTGAGGACGGAAGACGGTGCAGAAGAA
GCACTGAGAGTCCCAATGACCTGTCTTTGTGATCGATGGATCCAAGGATCTTGGAGAAGGAAATTTTGTAGGTCG
TGAAGCAGTTTGTCACTGGAATTTATAGATTCCTTGAACAATTTCCCCCAAGCCGCTCGAGTGGGGCTGCTCCAGT
ATTCCACACAGTCCACACAGATTCATCTGTAGAAACTTCAACTCAGCCAAAGACATGAAAAAGCCGTGGCCC
ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAAGAGAAATTTTACCCAG
GAGAAGGGGCGAGGCCCTTCCACAAGGGTGGCCAGAGCAGCATTTGTGTTCCACGACGAGCGGGCTCAGGATG
ACGTCTCCGAGTGGGCGAGTAAAGCAAGGCCAATGGTATCACTATGTATGCTGTGGGGTAGGAAAGCCATTG
AGGAGGAACTACAAGAGATTGCCCTGTAGCCCCACAAACAGCATCTCTTCTATGCCGAAGACTTCAGCACAAATGG
ATGAGATTAAGTAAAAACTCAAGAAAGGACATCTGTGAAGCTCTAGAGAGCTCCGATGGAAGACAGGACTCTCCAG
CAGGGGAACTGCCAAAAACGGTCCAAACAGCAACAGAATCTGAGCCAGTACCATAAATATCCAAGACCTATCTTT
CCTGTTCAAAAAACCTTCAAGAGGCCCTTTGGAAGAAAAACACGACTCAATGCAATGTGAAAAACCTTATAATGT
TCCAGAACCTTGC AAACGAAGAAGTAAAGAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAAAAATCGCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCACTTTATACCGGATTCAAT
GAACCGAGTGCAGAGGCCCAAGCTCAGGCTATTGTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGA
GAAACCTGTTTGGCCACAGCAAAAGACAAGAATATACACTAATCTGTATAAATTTATCTAGGAAAAAATCCCT
TCAGAATCTTAAGATGAATTTACAGGTGAGAATGAATAAGCTATGCAAGATTTTGTAAATATCTGTGGACAC
AATCTGCTTCTGCGCTCATCTGCTTGTAGTGTCAATCTCATTTGACTATGCAATAAGTTTGCACAAAGCTTACTT
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTGTACTGGAATTTACCTTGATATATGTATATGGATGTATG
CATAAAAATCATAGGACATATGTACTTTGTGGAACAGTTGGATTTTTTATACAAATATAAAATTCACCACTTCAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 15

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPTQTALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KRMRLHSTGTMGTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPDQSDVAEVAAKARD
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTLSVFQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
LAEDGKRCVADVYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCRINYCALNKPGC
EHECVNMEESSYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFII
NEDLKTCSRVDYCLLSDHGCYSCVNMDRSFACQCEGHVLRSDGKTCAKLDSICALGDHGCE
HSCVSSSEDSFVCQCFEGYILREDGKTCRRKDVQCAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA
HMKYMGKSGMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGKAIBBELQBIASEPTNKHLYAEDFSTMDBEISEKLKGGICEALEDSDGRQDS
PAGELPKTVQQPTSESPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL
BEKHDQCKCENLIMFQNLANEVVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,
781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,
639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,
540-546, 581-587

FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAAGTCCGTGCGTCCCGCGGCTGGCGCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCTGGTGACCAACTGTGTCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCAGCTGCAACTGCACTTGCCCGCCAACCGGTGTCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCAGCGTGGTACACCTTGACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGCAAAACGTGACCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGACCAGCATTAGATGTTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCCTGGAGCTGC
AGTGGTTGTCTGGAGCTGTTGTGGGTACCTGGTTGGACTGGGGTTGCTGGCTGGGTGGTCC
TCTTGTAACACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAAATGATATCAAGGAGGATGCC
ATTGCTCCCGGACCTTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAAATGGGACCTT
TTCTCTGTGTAACCTCCGACAGAGCCCTCCGGCCACCCCATGGGCCCTCCAGGCCCTGGTGCAT
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCACCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCTCTGGCTTGAGCCG
CATGGGTGCTGTGCTGTGATGGTGCTGCCAGAGTCAAGCTGGCTCTCTGGTATGATGAC
CCCACCACTCATTGGCTAAAGGATTGGGGTCTCTCCTTCCTATAAGGGTCACTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCAGTCACTGAGTCTCCAGGC
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTGAGGCTGGCTGGTTAGGTTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAATAAATAACATGAAATATGTTGTTTTCATTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 17

MISLPGLVLTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVFFVMWFFKQKEKEDQVLSYINGVTTSPKGVSLVYSMPSRNLSLRLEGLQEKDSGPY
SCSVNVQDKQGKSRGHSIKTLELNLVPPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQ
WDRQLPSFQTFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA
VVAGAVVGTLLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL
SSVTSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 18

CGCCACCCTGCGGCCACCGCCAAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTCCACTTTG
TTGAATTGTTTCTATCTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA
AATACGCAATGGAAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTACAA
TTTGTGAAGATGATAATGAATGTGGAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC
ACTAACACAGAAGGAAGTTATTATTGTATGTGTACCTGGCTTCAGATCCAGCAGTAACCA
AGACAGGTTTATCATAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCGCAT
TAGATAATGTCTGTATAGCTGCAAAATATTAATAAACTTTAAACAAAATCAGATCCATAAAA
GAACCTTGCGCTTTGCTACAAGAGTCTATAGAAATCTGTGACAGATCTTTCCACACAGA
TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTTACAGAAACA
CTATCTCAGCCAAGGACACCCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT
AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAAAC
ACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCAGAGCTTCC
AAAAGACCACAGAGTTTGATACAAATTTCAACGGATATAGCTCTCAAAGTTTTCTTTTGTAT
TCATATAACATGAAACATATTTCTCATATGAATATGGATGGAGACTACATAAATATATT
TCCAAAGAGAAAAGTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTATATTATA
AGAGTATTGTCCTTTGCTTTCATCTGACAACCTCTTATTGAAACCTCAAAATATTGAT
AATTCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTCACTGCTCAATGAGCTCAAAACC
ACCCACATTATGAACCTTGAAGAAAATAACATTTACATTAAAGTCATCGAAAGGTCACAGATA
GGTATAGGAGTCTATGTGCATTTTGGAAATTACTCACTGATACCATGAATGGCAGCTGCTCT
TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGTGTAATCACCT
GACACATTTTGAATTTTGATGTCTCTGCTGCTTCCATTGGTATTAAAGATTATAATATTC
TTACAAGGATCACTCAACTAGGAATAAATATTCTACTGATTTGTCTTGGCATATGCATTTT
ACCTTCTGGTCTCTTCAGTGAATTCAAAGCACCAGGACAACAATTCACAAAATCTTTGCTG
TAGCCTATTTCTGTGTAACCTTGTCTTTCTGTTGGGATCAATACAAATACTAATAAGCTCT
TCTGTTCAATCATTGCGGAGTCTACACTACTCTTTTGTAGTCTGTTTGCATGGATGTGC
ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTGTGA
CAAGAAATTTTATATCTTTGGCTATCTAAGCCAGCCGGTGTAGTTGGATTTTCGGCAGCAC
TAGGATACAGATATATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACCTTTATT
TGGAGTTTATAGGACCAGCATGCCTAATCATCTCTGTTAATCTCTTGGCTTTTGGAGTCAT
CATATACAAAGTTTTCTGTCACACTGCAGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACA
TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCTCTCGGCACCACCTGGATCTTT
GGGTTCTCCATGTTGTGCACGCATCATGGTTACAGCTTACCTCTTCACAGTCAGCAATGC
TTTCCAGGGGATGTTTCATTTTATCTCTGTGTGTTTATCTAGAAAAGATTCAAGAAGAAT
ATTACAGATTTGTTCAAAAATGTCCCTGTTGTTTGGATGTTTAAAGGTAACATAGAGAAATG
GTGGATAAATCAACTGCACAAAAATAAAATTCGAAGCTGTGGATGACCAATGTATAAAAA
TGACTCATCAAATATCCAATATTAACTACTAGACAAAAGTATTTTAAATCAGTTTTTCT
GTTTATGCTATAGGAATCTGATAGATAAATAAGTAAAATTTATGTATCATATAGATATACATGT
TTTTCTATGTGAATAGTTCTGTCAAAAATAGTATTGCAGATTTTGGAAAGTAATGGTTT
CTCAGGAGTGATATCATGCACCCAGGAAAGATTTTCTTTAAACACGAGAAGTATATGAA
TGTCCTGAAGAAATCACTGGCTGTATATTCTGTGACTGTGCTTTGAACTAGTCC
CCTACCACCTCGGTAAATGAGCTCCATTACAGAAAGTGAACATAAGAGAAATGAAGGGGAGA
ATATCAAAACAGTGAAAAGGGAAATGATAAGATGTATTTTGAATGAACCTGTTTTCTGTGAGAC
TAGCTGAGAAATGTTGTGACATAAAAATAAGAAATGAAGAAACACATTTTACCATTTTGTGAA
TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGTCTAAATCTGTTTCTT
TTTCTAATATTTCAAAAAAAAAAAAAAGGTTTACCTCCACAATTGAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 19

MKRLPLLVPFSTLLNCSYTONCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC
 GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENTVNANCHLDNV CIAA
 NINKTLTKIRSIKEPVALQEVYRNSVTDLSPTDIIITYIEILAESSSSLGYKNNTISAKDTL
 SNSTLTLEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLT KLMHTVEQATLRISQSFOKTTEFDT
 NSTDIALKVFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGP LLS
 SSDNFLLPKQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDYRSLCAF
 WNYSPDTMNGSWSSSEGCELTYSNETHTSCRCNHLTHFAILMSSGSPSIGIKDYNILTRITQLG
 IISLCLAI CIFTWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSSIAGL
 LHYFFLAFAWMCI EGIHLYLIVVGVYINKGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGT
 TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRTAGLKPEVSCFENIRSCARGA
 LALLFLGLTTWIFGVLVHVASVVTAYLFTVSNAPQGMFIFLFLCVLSRKIQEEYYRLFKNV
 PCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,
648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,
188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,
346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,
434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

[illegible]

TGGAACATATCTCTCCATATGAATATGGATGGAGACTACATAAATATATTTCCAAGNG
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATAAGAGTAT
TGGTCCCTTTGCTTTCATCATCTGACAACITCTTATTGAAACCTCAAATATATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTTCAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACCTTGAAAAATAACATTTACATTAAGTCATCGAAAGTCCAGATAGGTATA
GGAGTCTATGTGGCATTTTGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA
TTTTGCAATTTTGATGTCTCTGGTCCCTTCCATTTGGIATTAAGATTATAATATTTCTTACAA
GGATCACTCAACTAGGAATAATTATTCTCAGTATTGTTCTTGCCATATGCATTTTTTACCTTC
TGGTTCCTTCAGTGAATTTCAAGCACCCAGGA

FIGURE 21

GCTCCCAGCCAAGAACCTCGGGGCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
 CTAAGCGAGGCTCTCTCTCCCGCAGATCCGAACGGCCCTGGGCGGGGTACCCCCGGCTGGGA
 CAAGAAGCCGCGCCTGCGTCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
 GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
 TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG
 CCGCCGCGCCGTGAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
 GGCACTCACAGCCCCGAGCGCATCCCGGTGCGCGCCAGCCCTCCCGCACCCCCATCGCCGG
 AGTGTCCCGCAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGTCCACGTATGG
 ATCCTGGCCGGCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
 CCACGTGCACTACGGCTGGGGCGACCCCATCCGCTGCGGCACCTGTACACCTCCGGCCCCC
 ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGGCGTGTGGACTGCGCGCGGGGC
 CAGAGCGCGCACAGTTTGTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
 CGTGACACAGCTGCGGTACCTCTGCATGGGCGCCAGCGCAAGATGCAAGGGCTGCTTCACT
 ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCGCCAGATGGCTACAATGTGTACCGA
 TCCGAGAAGCACCGCCTCCCGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
 CAGAGGCTTTCTTCCACTCTCTCATTTCTGCCCCATGCTGCCATGGTCCAGAGGAGCCTG
 AGGACCTCAGGGGGCACTTGGAACTGACATGTTCTCTTCGCCCTGGAGACCGACAGCATG
 GACCCATTGTGGGCTTGTCAACGGAGTGGAGGCCGTGAGGAGTCCAGCTTTGAGAAGTAACT
 GAGACCATGCCCGGGCTCTTCACTGCTGCCAGGGCTGTGGTACCTGCAGCGTGGGGGACG
 TGCTTCTACAAGAACAGTCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA
 GTTGTACATATTCAAGATTTTCCATTGGCAGTGCCAGTTTCTAGCCAAATAGACTTGTCTGAT
 CATAACATTGTAAGCTGTAGCTTGCCAGCTGCTGCCGGGGCCCCATTCTGCTCCCTCGA
 GGTGTCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
 TCACCTTCTTTGAAAAATTCTTATGTCAAGCTGAAATCTCTAATTTTTTCTCATCACTTC
 CCCAGGAGCAGCCAGAAGCAGGCAGTAGTTTAAATTTAGGAACAGGTGATCCACTCTGTA
 AAACAGCAGGTAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
 ACCATTTGCCCTTCCCAAATCCCTCAGGCCAGAATGACTGGAGCAGGCATGGCCACCAG
 GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACCTTGAGAATTCCTC
 CTGAGGCCAGTTCTGTCTATGGATGCTGTCTGAGAATAAATCTGCTGTCCCGGTGTCACTGTC
 TTCCATCTCCCAGCCACAGCCCTCTGCCACCTCACATGCCTCCCATGGATTGGGGCCT
 CCCAGGCCCCCACTTATGTCAACCTGCACCTTCTGTGTTCAAAAATCAGGAAAAGAAAAGAT
 TTGAAGACCCCAAGTCTTGTCAATAACTTGTGTGTGGAAGCAGCGGGGAAGACCTAGAAC
 CCTTCCCCAGCACTTGGTTTTTCAACATGATATTTATGAGTAATTTTGTGATATGTACA
 TCTCTTATTTTCTTACATTATTTATGCCCCAAATATATTTATGTATGTAAGTGAGGTTG
 TTTTGTATATATAAATGGAGTTTGTTTGT

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 22

MRSGCVVHVWILAGLWLA VAGRPLAFSDAGPHVHYGWDPIRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVTGLEAVRSPSF EK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCTCCTGCGCTCCTGCCGCGGGACCCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCCGGGAAGATGCGCGAGGAGGAGCCGCCACCGCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTGCCTGGGCTATCATAAGGCCATATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAACTGGGTGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAACTGGAACT
CTGCAATTTAATACTGTTTCCAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTGGATATCGCAGGTGTCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCGGTTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTCCAGAAGAGTAATCTTCATCTAA
AGCCACGACAATGAGTGAAAAATGTGCAGTGGCTCACGCCTGTAATCCAGCACTTTGGAAGG
CCGCGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCACTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCTGCAGTTCCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAAACAGAGCAAGATTCATCTCAAAAAATAAAATAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTTTACAATAAATATAGCTTGATATTC

FIGURE 24

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEBAILACKTPKKTVSSRLE
WKKLGRSVSFVYYQQLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQQNLEED
TVTLEVLVAPAVPSCSEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST
NSSYTMNTKTGTQLQFNTVSKLDTGEYSCEARNVSGYRRCFPGKRMQVDDLNISGIIAAVVVVA
LVISVCGLGVCYAQRKGYFSKETSFOKSNSSSKATTMSENVQWLTPVIPALWKAAGGSRGQEF

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
 AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCCGAGCCCTGGCAT
 CATGCTGCTATTCTCGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA
 ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA
 TTATATCATTAAAGGAAATGTAACCTTCTCTTCTCCAATATGCATGACATTTTTGGACAATG
 CAATTGGGGCACTGGCACTTATTTTCAGTGAAGAAAAAATTTGGGTTCTATGGCATTCATCA
 TTTGACAAATGCAAGCATCTTCTTATCAATCAGCTCCTATTGAACCTTACTAGCACTGACTG
 TGAATCCTTAAGGGCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
 CCGAATTCATGTGCTACTTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
 TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCGCTTGGTTTACACCCAGATCCATTAT
 ATGGAAGCATCTACAGTGGATTGTAATGATTAGGTCTTTTAACTTTCCAGCCAGATTGCC
 AGCTAACACACAGATTCTTCTCTACAGACTAACAAATATTGCAAAAATGGAATACTCCACAG
 ACTTTCCAGTAAACCTTACTGGCCTGGATTATCTCAAAACAATTTATCTTCAGTCACCAAT
 ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAAACAACCTTACTGA
 ACTGCCTGAAAAATGTCTGTCCGAACCTGAGCAACTTACAAGAACTCTATATTAATCACAAC
 TGCTTTCTACAATTTCACTGGAGCCTTTATTGGCCTACATAATCTCTTTCGACTTCATCTC
 AATTCAAGATGATGACAGATGATCAACAGTAGTGGTTTGTAGTCTCTTCCAATCTAGAGAT
 TCTGATGATTGGGCAAAATCCAATTATCAGAATCAAAGACATGAACCTTAAAGCCTCTTATCA
 ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT
 GGACTGAAAACTTAGAAAGCATCTCTTTTACGATAACAGGCTTATTAAGTACCCCATGT
 TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTGGATCTAAATAAAAATCCTATTAATAGAA
 TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAGAGTTGGGGATAAATAATATGCCT
 GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
 TACTAACACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG
 AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCTGTACCATGGTACCATTTGAGTCTCTG
 CCAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCACTCGTTG
 GATGAACATGAACAAAACCAACATTCGATTTCAGGAGCCAGATTCACTGTTTTGCTGGTGAAC
 CACCTGAATTCGAAGGTCAAGATGTTCCGGCAAGTGCATTTACGGGACATGATGGAATTTGT
 CTCCTCTTATAGCTCCTGAGAGCTTTCTCTTCTAAATCTAAATGTAGAAGCTGGGAGCTATGT
 TTCTTTTCACTGTAGAGCTACTGCAGAACCACAGCTGAAATCTACTGGATAACACCTTCTG
 GTCAAAAACCTTTCCTAATACCTGACAGACAAGTTCTATGTCATTCAGGGAACACTA
 GATATAAATGGCGTAACTCCCAAAGAGGGGTTTATATACCTGTATAGCACTAACCTTAGT
 TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG
 GCTCTTTGAATATTAATAAAGAGATATTCAGGCCAATTCAGTTTGGTGTCTGGAAAGCA
 AGTTCTAAAATCTCAAATCTAGTGTAAATGGACAGCCTTTGTCAAGAGTGAATAATCTCA
 TGTGCGCAAGTGTCTGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC
 CATCAACTGAGTATAAAATTTGTATTGATATTTCCACCATCTATCAGAAAAACAGAAAAAA
 TGTGTAAGTGTCAACCAAGGGTTTGCACCTGATCAAAAAGAGATTGAAAAGATAATATAC
 CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGCTTTATCA
 GCTGCTCTCTCCCAAGAAATGAACTGTGATGGTGGACAAGCTATGTGAGGAATTTACTACAG
 AAACCAACCTTTGCATTAGGTGAGCTTTATCTCTCTGTATAAATCTCTGGGAAGCAGGAAA
 AGAAAAAGTACATCACTGAAAGTAAAGCAACTGTTATAGGTTTACCAACAAATATGTCCCT
AAAAACCAAGCAAGCAACTACTCCAAAATGAAC

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT
FPARLPANTQIILLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTPELPEKCLSELNLQELYINHNLLSTISPGAFIGLHNLRLHLNSNRLQMINSKWFDA
LPNLEILMIGENPIIRIKDMNFKPLINLRSVIAGINLTEIPDNALVGLENLESISFYDNRL
IKVPHVALQKVNNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLDP
LRKIEATNNPRLSYIHPNAFFRLPKLESMLNSNALSAHYHGTIESLPNLKEISHSNPIRC
DCVIRWMNMNKTNIREFMEPDSLFCVDPPEFQGNVRQVHFRDMMEICPLIAPESFPSNLNV
EAGSYVSFHCRTAEFPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRIDIQANSVLVSWKASSKILKSSVKWTAfV
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTKGLHPDQKE
YEKNNTTTLMACLGGLLGIIGVICLISCLSPENMCDGGHSYVRNYLQKPTFALGELYPPLIN
LWEAGKEKSTSLVKATVIGLPTNMS

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

FIGURE 27

GCCCCGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA
TGATACTGTGCTTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGCTTTGTCTCTCCTCTGGG
GGTTTTAAATGTCACCTGTAGCAATGCAAAATCTCAAGGAAATACCTAGAGATCTTCTCTCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCAC
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCTCTCAATGCTGC
CAACGACGCTGACCTTTGTAACTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG
GATGCCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATTAGTGTCCAACTGACTGTTCATTGAGAAAGAAAGAAA
GTAGTTTGCATTGCAGTAGAAATAAGTGGTTTTACTTCTCCATCCATTGTAAACATTTGAA
ACTTTGTATTTCAGTTTTTTTTTGAATTATGCCACTGCTGAACTTTTAAACAACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAAGCAAATAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAACA

FIGURE 28

MNLVDLWLTRSLSMCLLLQS FVLMILCFHSASMC PKGCLCSSSGGLNVTCSNANLKEIPRDL
PPETVLLYLDNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGV AETLQTLDLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVI CKTSVLDEHAGRPFLL
NAANDADLCNLPKKTDDYAMLVTMFGWFTMVISYVVVYVRQNQEDARRHLEYLKS LPSRQKK
ADEPDDISTVV

Signal sequence:

amino acids 1-33

Transmembrane domain:

amino acids 205-220

N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

FIGURE 29

ACCAGCGCGAGCGGACCGAAGGCGCGCCCGAGATGCGAGGTGAGCAAGAGGATGCTGGCGGGG
GGCGTGAGGAGCATGCCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT
GGGCTCAGTGTGTGTCAGGCTCGGCCACGGGGTGCCGCCCCCGTGCAGTGTCTCCGCCAGG
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG
ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCCGCCAG
CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAAACATCGTGAGCGCCGTGGAGCCCGCG
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAAACAAGAT
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
ACAATGACCTCGTCTACATCTCTCACCAGCGCTTTCAGCGCCCTCAACAGCCTGGAGCAGCTG
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACGAGGCGCTGTCCCACTGCACGGCTT
CATCGTCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAAC
TGCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCCGCTGTGCC
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCTCAACCTCTCTTACAACCCCATCA
GCACATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCAGCTGGTGGGC
GGGCAGCTGGCCGTGGTGGAGCCCTATGCCCTTCCGCGGCTCAACTACCTGCGCGTGTCAA
TGTCTCTGGCAACCACTGACCACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
AGACACTCATCTGGACTCCAACCCGCTGGCCTGCGACTGTGCGCTCTGTGGGTGTTCCGG
CGCCGCTGGCGGCTCAACTTCAACCCGCGAGCAGCCACGTCGCGCCACGCGCCGAGTTTGTC
GGCAAGAGGTTCAAGGACTTCCCTGATGTGCTACTGCCCACTACTTCACTGCGCGCGCG
CCCCCATCCGGGACCGCAAGGCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAAGTTT
GTGTGCGGGCGGATGGCGACCCGCGCCGCCATCCTCTGGCTCTACCCCGAAGACCTT
GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCAGCTGGAGGTGCGCT
ACGCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACCGGGCGGCAACGAC
TCCATGCCCGCCACCTGCATGTGCGCAGCTACTCGCCGACTGGCCCCATCAGCCCAACAA
GACCTTCGCTTTTATCTCCAACAGCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG
TGCCTTTCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC
CTGGGCGTCGTCTCTTTCGCTGGTGTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC
AAAGCACAAACATCGAGATCGAGTATGTGCCCGAAAGTCGAGCGCAGGCATCAGTCCGCG
ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCGCGGGCGGGGGGAGGGACCCCG
GGCGCGGGCAGGGGAAGGGGCTGCTGCCACCTGCTCACTCTCAAGTCTTCCACCTC
CTCCCTACCTTCTACACACGTTCTCTTCTCCCTCCCGCTCCGTCCCTGCTGCCCGCG
CCAGCCCTCACCACCTGCCCTCTTCTACAGGACCTCAGAAGCCAGACCTGGGGACCCCA
CCTACAGGGGCTTGAAGACTGGAGTTGAAAGCCGACCAACCGACACGCGCAGAGTCA
ATAATTCAATAAAAAAGTTACGAACCTTCTCTGTAACCTTGGGTTTCAATAATTATGGATTTT
TATGAAAACCTTGAATAATAAAAAGAGAAAAAACTAAAAA

FIGURE 30

MQVSKRMLAGGVRSMPSPLLACWQPIILLVLGSLVSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLLDGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFNNLNLRLTL
GLRSNRKLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA
FSGLSNLEQLTLEKCNLTSTIPTEALSHLHGLIVLRLRHLNINAIIRDYSFKRLYRLKVLEISH
WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLYLRLFNLSYNPISTIEGSMLEL
LRLQEIQLVGGQLAVVEPYAFRGLNYLRLVNLVSGNQLTTLEESVFHSVGNLETLILDSNPLA
CDRLRLWVFRRRWRNLNFRQQPTCATPEFVQGEFKDFPDVLLPNYFTCRRARIIDRKAQV
FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGLRTVFPDGTLEVRYAQVQDNGTYL
CIAANAGGNDMPAHLHVRYSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISPLGVVLFCLVLLFLWSRGKGNTKHNIIEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,
492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,
607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGCGCCCTTTCGGTCAAC
ATCGTAGTCCACCCCCCTCCCATCCCCAGCCCCGGGGATTTCAGGCTCGCCAGCGCCAGCC
AGGGAGCCGGCCGGGAAGCGCGATGGGGGCCCCAGCCGCTCGCTCCTGCTCCTGCTCCTGCTGC
TGTTTCGCTGTGCTGGGCGCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGTCAAGTGC CAAGTGAAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA
GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGAAGCCCATCATCACTGGTT
ATAAATCTTCAATACGGGAAAAAGACACAGCCACCCTAAACTGTCACTCTTCTGGGAGCAAG
CTGCGAGCCCGGCTCACCTGGAGAAAGGGTGACCAAGAAGTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCCAATGTTAAACCTTCACTGTGACGAGCTCGGTGACATTCCAGGTTACCC
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT
GCCCTGATCTTCCCTTTCTCAACAAGAGTGACAGTGGCACCTACGGTGCACAGCCACCAG
CAACATGGGCAGCTACAAGGCCTACTACACCTCAATGTTAATGACCCAGTCCGGTGCCTT
CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCTGCTG
CTCATCATGCTCATCTTCCTTGCCACTACTTGATCCGCGACAAAGGAACCTACCTGACACA
TGAGGCCAAAAGGCTCCGAGCATGCTCCAGACGCGACACGGCCATCATCAATGCAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTTCATCTTAGAGGCGCCTGCCCACTTCTGTC
GCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTACCAACCCGGCACTTGTACAGAGCAA
CCGCGAGGGCCGCCCTCCCGCTTGGTCCCCAGCCACCCACCCCTGTACAGAATGTCTGTC
TTTGGGTGCGGTTTGTACTCGGTTTGAATGGGGAGGGAGGGCGGGGGAGGGGAGGG
TTGCCCTCAGCCCTTTCCTGGCTTCTCTGCATTTGGGTTATTATTATTTTGTAAACAATCC
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACA
AACAAAAACA

FIGURE 32

MGAPAASLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLCQVKDHEDSSLQW
SNPAQQTLYFGEKRALRDNRILVLTSTPHELISISNVALADEGEYTCSTFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEFTRIQUEDPNGK
TFTVSSSVTTFQVTRREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ
KLLHLHCEGRGNFVPQQYLWEKEGSVPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA
YYTLNVNDPSPVPSSSSTYHAIIIGGIVAFIVFLLIMLIFLGHYLRHKGTYLTHEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCCTTTTCTTCTCCTTTCTCTGG
 CTTTCGGACATTTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
 TTACTTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT
 GCTGGAGACGCTCTCTTTGTTTTGCCGTGGAACGTTACAGGGGACGTTTTCGAAAGAGAAGA
 TCTGTTTCTGCAATGAGATAGAAAGGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA
 AGTCTGCAGCGCTTTCACTGCCCGGACTTCCGAGTTTACCATTATTCTGTCATGGCAATTC
 CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG
 AAAACAATGGCTTGCATGAAATCGTTCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAGAAG
 CTGCACATCAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
 TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG
 ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCTACCTGCCAAC
 GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
 CTATGAGGAGGTCTTGGAGCAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCCT
 GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAAACATTCCTCAAGAATGCC
 CTGATCGGCCGAGTGGTCTGCGAAGCCCCACAGACTGCAGGGTAAAGACCTCAATGAAC
 CACCGAACAGGACTTGTGCTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCCCCCTG
 CCCAAGAGAGACCTTTGCTCCTGGACCCCTGCCAATCTCTTTCAAGACAATGGGCAAGAG
 GATCATGCCACACCGAGGTCTGCTCCTCAACCGAGGTACAAAGATCCAGGCAACTGGGAGAT
 CAAAATCAGACCCACAGCAGCGATAGCGACGGTAGCTCCAGGAACAAACCTTAGCTAAC
 GTTTACCTGCCCTGGGGGCTGCAGCTGCGACCACATCCAGGGTCGGGTTTAAAGATGAAC
 TGCAACAAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTTAACGTGACGGA
 GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA
 ACCTCATTTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACTTTCAAG
 AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACAGCGTCTCCCGGA
 GAAATTCGGGGGCTGCAAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGTCA
 TCCTCCCGGGGCACTTTCAATGCCATGCCCAAACCTGAGGATCCTCATTTCTCAACAACAACCTG
 CTGAGGTCCCTGCTGCTGTGGACGTGTTGCTGGGGTCTGCTCTCTAAACTCAGCCTGCACAA
 CAATTACTTATGTACCTCCCGGTGGCAGGGGTGCTGGACAGTTAACCTCCATCATCCAGA
 TAGACCTCCACGGAACCCCTGGGAGTGTCTCTGCACAAATTGTGCTTTCAAGCAGTGGGCA
 GAACGCTTGGGTTCCGAAGTGTGATGAGCGACCTCAAGTGTGAGACGCCGTGAACTTCTT
 TAGAAAGGATTTCTAGCTCTCTCCTCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
 CGCCCAAGTTAACTTTCGACAGTAAACACGACACTGGGTTGGCGGAGACGGAGCGCATCC
 AACTCTACCTAGACACCAGCAGGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT
 GTTTGTCACTCCCGCTTCCACCGTGGTGGGATGCTCGTGTATTCTGAGGAACCGAAAGC
 GGTCCAAGACGACGAGATGCCAATCTCTCCGCGTCCGAGATTAATTCCTACAGACAGTCTGT
 GACTCTTCTCTAGTGCACAAATGGGCCTTACAACGAGATGGGGCCACAGAGTGTATGACTG
 TGGCTCTCACTCGCTCTCAGACTAAGACCCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG
 ATACATCTTCCCCACCGCAGGACCCCGGGGCTGGAGGGGCGTGTACCCAAATCCCGCG
 CCATCAGCCTGGATGGGCAATAAGTAGATAAATAACTGTGAGCTTCGCACAACCGAAGGGGCT
 GACCCCTTACTTACTCTCCCTCTTGAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
 GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTTGACAGAAAGCCAGCAGCACCCTGCTGGAAG
 AACTGACAGTGCCTCGCCCTCGGCCCGGGGCTGTGGGGTTGGATGCGCGGTTCTATAC
 ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTCCCTGTGGATTAG
 CCCCGTGATGCTCCTGTGGTACGCGAGGATGGGCAGTTGCACGAAGGCATGAATGTAT
 TGTAAATAAGTAACCTTTGACTTCTGAC

FIGURE 34

MLLWILLETSCLCFAAGNVTGDVCKEIKICSCNEIEGDLHVDCEKKGFTSLQRFTAPT SQFYH
LFLHGNSLTRLPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNKKIKSFRKQ
TFLGLDDLEYLQADFNLLRDI DPAGAFQDLNKLEVLILNDNLISTLPANVFQYVPI THLDLRG
NRLKTLTPYEEVLEQIPGIABILL EDNPWDCTCDLLSLKLEWLENI PKNALIGRVVCEAPTRLQ
GKDLNTEETQDLCPLKNRVDSLSLPAAPAEETFA PGPLPTPFKTINGQEDHATPGSAPNGGTK
IPGNWQIKIRPTAAIATGSSRNKPLANS LCPGGCSCDHIPGSGLKMN CNNRNVSSLADLKP
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLDDLRLWLYMDSNY
LDTLSREKFAGLQNL EYLNVEYNAIQ LILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC
ETPVNFPRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL
VPGLLLVFVTSFTVVGMLVFILRNKRKSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG
AHRVYDCGSHSLSD

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 337-348

FIGURE 35

AGTCGACTGCGTCCCTGTACCCGGCGCCAGCTGTGTTCTTGACCCAGAACTAACTCAGGGC
 TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCTGTGCGGGCCTAAGGGAACTGTTGGC
 CGCTGGGCCCCGCGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG
 AAGGGAGGGGAAACCGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
 AGCTCTGCGTCTCTGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGGGGGCTCAG
 AGAATGAGGCGCGGCTTCGCCCTGTGCCTCTCTGGCAGGCGCTCTGGCCCCGGGCGGGCGG
 CGGCGAACACCCCACTGCGCAGCGTGTGCTGCTGCGGCTCTGGGGGCTGTCTACAGGCTTCG
 ACCACGCTACCATGAAGCGCGCAGGCGGCGGAGGCGCTGCATCTGTGAGGTGGGGCGCTC
 AGCACGCTGCGTGGGGCGCGAGCTGCGCGCTGTGCTGCGGCTCTGCGGGCAGGCGCCAGG
 GCCCGGAGGGGGCTC CAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCCACT
 GCACCTGGAGAACGAGCCTTTGCGGGGTTTCTCTGGCTGTCCTCCGACCCCGGCGGTCTC
 GAAAGCGACACGCTG CAGTGGGTGGAGGAGCCCCAACGCTCTG CACCGCGCGAGATGCGC
 GGTACTCCAGGCCACCGGTGGGGTGGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC
 GCGCCAACGCTACCTGTGCAAGTACAGTTTGGGTCTTGTGTCTGCGCGCGCGCCCCGGG
 GCGGCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCGCTCTGGACTTCAG
 TCCACTGGGACCGAGGTGAGTGCCTCTGCGGGGACAGCTCCCGATCTCAGTTACTTGCAT
 TCGCGAGCAAAATCGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC
 GGGAGGTACTCTCGGTGCTGGCAATGCGCAGAGCTCCCTAACTGCGCTAGACGACTTGGGAGG
 CTTTGCTTGCAGATGTGCTACGGGCTTCGAGCTGGGGAAGGACGCGCGCTCTTGTGTGACCA
 GTGGGAAGGACAGCCGACCCTTGGGGGACCGGGGTGCCCCACAGGCGCCCCGCGGCACT
 GCAACCAGCCCCGTGCGCAGAGAAATGCGCAATCAGGGTCGACGAGAAGCTGGGAGAGAC
 ACCCATTTGCCCTGAACAAGACAATTAGTAACATCTATTTCTGAGATTTCTCGATGGGGAT
 CACAGAGCAGATGTCTACCTTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC
 CCATCAGGAGCGTGTATTTCAAAGTTTAATTTCTACGACTTCTCTGCCACTCTCAGGCTTT
 CGACTCTCTCTGCGCGTGGTCTTCATATTTGTGAGCAGCAGTAGTAGTGTGGTGATCT
 TGACCATGACAGTACTGGGGCTGTCAAGCTCTGCTTTACGAAAGCCCCCTCTTCCAGCCA
 AGGAAGGAGTCTATGGGCCCGCGGGCCTGGAGAGTGATCCTGAGCCCGTGCTTTGGGCTC
 CAGTTCTGCACATTGCACAAA CAATGGGGTGAAAGTCCGGGACTGTGATCTGCGGGACAGAG
 CAGAGGGTGCTTGTGCGGAGTCCCTCTTGGCTCTAGTGATGCATAGGGAACAGGGGA
 CATGGGCACTCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAATTAC
 TTTGTGTAATGACAATTTCTGCAGAAATCCCCCTTCTCTAAATTTCCCTTACTCCACTGAG
 GAGCTAAATCAGACTGCACATCTCTCCCTGATGATAGAGGAAGTGGAAAGTGCCTTTAGGA
 TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA
 TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAAACAAATAGAACACAATAAATTACA
 TTAAAAATAATTTCTACCAAATGGAAAGAAATGTTCTATGTTGTTGAGGCTAGGATAT
 ATTGGTTCGAATCCAGGGAAAAAATAAAAAATAAAAAATAAAGGATTGTTGAT

100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAE LRAVLALLRAGPGPGGGS KDLLFWVALERRRSHCTLENEPLRGFSWLS SDPGGLE
SDTLQWVEEPQORSTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGA
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDLVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPATA
TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILMTVLGLVKLCFHESPPSQPR
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

FIGURE 37

CGGACGCGTGGGATTACAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAAGTGCCCTCCGCCCTGCCGGCCGCGTATC
CCCCGGCTACCTGGGCCCGCCCGCGCGCGGTGCGCGCGTGAAGAGGAGCGCGCGGCGAGCCGA
GCGCCGGTGTGAGCGCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGGGTGCGGA
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAAACCCGAGCGTCTACGCTGCCATGA
GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG
CAGCAGTCCCCAGAGAGACCTGTTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
TATTGCGAGTGAAGTTTTTCCTGGAGTGTACCCCTCCAAATAGCAAAATGTACTTGGAAAATCA
CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTTCATAGACCTCGAGAGTGACAAC
CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG
CTTCTGTGGCACTTTCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAC
CCCCAAGTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC
CAAAGAATCAGCTTATAGAATTAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC
CGATATGATTATGTGGCTGTGTTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA
GTATTGTGGTGATAGTCCACCTGCGCCAATGTGTCTGAGAGAAATGAACCTTCTTATTCAGT
TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTACAGGCCAAAA
AAACTGCCTACAACCTACAGAACAGCCTGTCAACCACCAATTCCCTGTAACCCAGGGTTTAAA
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAAACCATCACTCGCGATGGGAGTTTG
CAGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTACAGCAGCGCGG
CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC
TTTATCATGATGTTCAAGACCAAGAAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG
TTAAACAGTGAACTGTGTCCATTAAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTT
TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCGAAAGATGG
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCCTCCGAGATAGCTGAGGGAAGTTCTT
TGCCCTGCTGTGAGAGGAGCAGCTATCTGATTGGAACCTGCGGACTTAGTGCGGTGATAGGA
AGCTAAAAGTGTC AAGCGTTGACAGCTTGAAGCGTTTTATTTATACATCTCTGTAAAAGGAT
ATTTTGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTGAAGTGCAATATTTATAGT
GTTATTGTGTTTCAAGCCTTGGCCCTGAGGTGTTACAATCTGTCTTGCCTTTTCTA
AATCAATGCTTAATAAAATATTTTTAAAGGAAAAA

FIGURE 38

MRGANAWAPLCLLAAATQLSRQQSPERPVFVTCGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVLNFRFIDLESNLCRYDFVDVYNHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAEPNERGDQYCGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAIVSERNELLI
QFLSDLSLTADGFIGHYIFRPKKLPITTTEQPVTTTFPVITGLKPTVALCQQKCRRTGTLEG
YCSSDFVLAGTVITITITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCACGCGCCCGCGGGCTGGGGCGGTGCGTTCTT
CCTTCTCCGTGGCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCTCAACCTCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCAGTCTTCTCCCCCGCCTCAGCCCCATCGTGTACATCTG
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACCTTTGGAG
GTGGAACACTGCCTGGGAGGAAGAGAATTGTCCAATACAAAGACAGTGAGACCCGCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGGCCTCTGTC
CTTCCCTGTCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT
TTTGGCCCCGTGCCCCGATGCTCAGGACCTGAGGAATCAAAGTGTTCGAATGCAAGAAGGG
CTGGGCCCTGCATCACTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAAGT
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTCGCTGTAAGAAGGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTCTCCGGGAGAGA
ACAAGCAGTGTGAAAAACCCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGTGTCAGCAGATGTCTTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTGAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGCAGATA
ATCGCGGCCACCACCTGTAGGACCTCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCC
TCCTGTGAGACACTCAGGACAGCTTGGTTTATTTTGGAGAGTGGGTAAAGCACCCCTACCTG
CCTTACAGAGCAGCCAGGTACCCAGGCCCGGGCAGACAAGGCCCTTGGGTAAAAAGTAGC
CCTGAAGGTGGATACCATGAGCTCTTACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA
ATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCTGCCAGCTGCATGCTGCCAGTTCTCTGT
TCTGTGTTCAACACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA
AAGTCTTGGAAGTTAAAAAAAAAAAAAAAAAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 40

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQSPPPQSSPPPPQHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKLCPPAGTFGPSCLPCEGGTERPCGGYGQCEGEGTRGGSGHCDQAG
YGGACQCCGLGYFAERNASHLVCSACFGPCARCSGPEESNLCQKKGWALHHLKCVDI
CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSCKCLDVDECE
TEVCPGENKQCENTEGGYRCICABGYKQMEGICVKEQIPESAGFFSEMTEDELVLVQLQMF
FGIIICALATLAAKGDVFTAFIFGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:

amino acids 1-29

Transmembrane domain:

amino acids 372-395

N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 290-294

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,
326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

FIGURE 41

TGAGACCTCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCTTGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCACCTTGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGCCCCAGT
ACGTGGCCCCGCTGCAGCGCAGCCACGGGGACCGTCCCGCGGAAAGAGGTTACGCCAGAGC
TTCGAGAGGTTGGCCGGCAGGTTTCCTGGCGTTGGAGGCCAGCACACCTGCTGGTGTTCGG
CATGGAGCAGCGGGTGC CGCCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGGTGTCCCCGCGCAGCGCCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCCTTCGACGTGACCGAGGCCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCCGGCAGCCGCTGCTGCTACAGGTGTGGTGCAGAGG
GAGCATCTGGGCCCGCTGGCGTCCGGCGCCACAAGCTGGTCCGCTTTTGCTTCGAGGGGGC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCGGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAAGTGGGTGCTGGAGCCCCGGGCTT
CCTGGCTTATGAGTGTGTGGGCACCTGCGGCAGCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAAGGAGGCTCCAGCCATTAGCGCCCTAGTG
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTTCGAGGGTACCAGGAGAGCTG
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT
CCTCTGACAAGTTACCTCACCTAATTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACCTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCCANTGTGTGCTATTGTTTACTTGTCTGTGAC
TGGATCTGGGCTAAAGTCCTCCACCACCACCTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAGACTTTGTAAAACATGAATAAAACATTTTATTCT
AAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFRVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRRLFQEP
VPKAAALHRHGRLSPR SARARVTVEWLRVRDDGSNRTSLIDSR LSVSHESGWKAFDVTEAVNF
WQQLSRPROPLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLDGYGAQ
GDCDPEAPMTGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCTCRQPPEALAFKWP
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVS LPNMRVQKSCASDGALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

FIGURE 43

GTCTGTTCCACAGGAGTCCTTCGGCGGCTGTTGTGTGTCAGTGGCCTGATCGCGATGGGGACAAA
 GCGCGAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
 CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCTTGAGAATAATCCT
 GTGAAGTTGTCTGTGCCTACTCGGGCTTTTCTTCTCCCGTGTGGAGTGGAAAGTTTGACCA
 AGGAGACACCACAGACTCGTTTGTCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
 TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
 ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
 GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGGAACCGGG
 CAGTGTGACATGCTCAGAACAAGATGGTTCGCCACCTTCTGAATACACCTGGTTCAAAGAT
 GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCTATGTCTT
 GAATCCCAACACAGGAGAGCTGGTCTTTGATCCCCGTGCAGCCTCTGATACTGGAGAATACA
 GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCCATGGAAGCT
 GTGGAGCGGAATGTGGGGGTATCGTGGCAGCCGTCTTGTAAACCTGATTCTCTGGGAAT
 CTTGGTTTTTGGCATCTGGTTTGCCCTATAGCCGAGGCCACTTTGACAGAACAAGAAAGGA
 CTTGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
 ACCTCGTCATTCTCGGTGTCAGCCTGGTTCGGCTCACCGCCTATCATCTGCATTGTGCCTTACT
 CAGGTGCTACCGGACTCTGGGCCCTGATGTCTGTAGTTTACAGGATGCCTTATTTGTCTTC
 TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
 ATCTCTCTTACATGCCCTCCCTCCCTTCTCTACCACTGCTGAGTGGCTGGAACCTGTTTAAA
 GTGTTTTATCCCCATTTCTTTGAGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
 TTCTAAGTAGACAGCAAAAATGGCGGGGTGCGAGGAATCTGCACTCAACTGCCACCTGGC
 TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTCCCTTGTGTACTGAC
 GACCAGGGCCAGCTGTTCTAGAGCGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
 TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
 CCACTGGGATCCCTCTGCCCCGTCTCTCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
 GGAAAAATGGGAGCTCTGTTGTGGAGAGCATAGTAAATTTTTCAGAGAACTTGAAGCCTAAAG
 GATTTAAAACCGCTGCTCTAAAGAAAAAGAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG
 TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGAGTTCCGGATCAGCCTGACCA
 ACATGGAGAAACCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGATGCCTGTAGTC
 CCAGCTGCTCAGGAGCCTGGCAACAAGACAAAACCTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTRTLVCYNNKITASYEDRVTFLLPTGITFKSVTREDTGTYTCMVSEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTGMTSNAVRMEAVERNVGVIVA AVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTCGGCCTCGGACTAGGCCT
GGAGGCCGCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGCCCCAGCTCAG
GCTCGTGCCACCCACCAAGTTCCAGTGCCGCACCACTGGCTTATGCGTGCCCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAAGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCACCGCCCCCTGGCCTCCCTGCCCCCTGCACCGGCGTCA
GTGACTGCTCTGGGGGAAGTGACAAGAACTGCGCAACTGCAGCCGCTGGCCTGCCTAGCA
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGCGCA
CCCAGACTGTCCCAGCTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCCTGGAGAGTGTCACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCCTGGAGAGTGTCCTCTGTGCGGAATGCCACATCCTC
CTCTGCCGAGACCACTCTGGAAGCCCACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTGTCTGGCTCCGAGCCCAGGAGCGC
CTCCGCCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC
CTCGCTGCCCCTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAAGTGA
GGGCTGGCCCCAGGCAGCTCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR
TSGLCVPLTWRCRDLDGSDGSEEEECRIEPTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL
RNC SRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSRLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPATAYGVIAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,
224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGAGGAGGCGGAGGAGGAAGGAGTCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCAGGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCGAGCG
GAGGGGTTTCTTTGAAGGCTCTGTAGCCCGATTCTACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAGAGACTGTGTTTGAAGCATTTTAAATGGAACCTTAGGCTGGATCCCAGTGTA
TAATCCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCTATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCGGGATTTAAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCCACCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTTCGTCT
GCCACCCGCGGCTTGTGAGCGCTACAACCACGGAACCTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCTTCT
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCAGCACCCATGAGACCTCCTGA
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGTCTGCTGCTGCTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTCCCCCAGGGGGCCTCCCCGGAG
TTCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCTCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCTTAAGAACTGATTGATTAAAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCTTTCTTCTCTTGGTTTTAGACAAATGTAAACAA
AGCTCTGACTCTTAAATGCTATGCTGATAGATGGTGGAGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAAATTAAGNAAAAA

FIGURE 48

MYHGMNPSNGDGFLEQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETLLTTWKIVAFATSVLLVLLLVILARMFQTKFKAHFPPRGPPRSSSSDPPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLFVDDQSPPAYPGSGD TDTGPGESETCDS
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWLFLRN

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,
364-368, 380-384, 467-471, 468-472

N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 394-405

FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCCGTGGCCCTAGAGA
 TGCTGCTGCCGCGGTTGCAGTTGTGCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCGT
 AGCGCCCGAGTGTGCGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG
 CCGTGCTGTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCC
 TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA
 TAAAGTCATTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT
 GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAACTGATAGAA
 AAGTTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGGCTCAGGAGGCGTGA
 GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGAATGATGCGAGCATAT
 CACAATTTAGGAACTGGTATGTGGATGAGCCGTCTGCCGCGAGCGAGGTCTGCGTGGTCATG
 TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAAATGATGA
 CCGGTGCAACATGAAGAAACATTTTCATTGCAAATATTCTGATGAGAAACCAGCAGTTCCCT
 CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG
 GAAGAAGATGCCAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCCTGAATCTGGCCATACAT
 CCTAATCCCCAGCATTTCCCTTCTCCTCCTTGTGGTGACCCACAGTTGTATGTTGGGTTT
 GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC
 TGGCCCTCTCCTCACCAGGGAAAACAGCCCGGACCTAGAGGTCTACAATGTGATAAGAAAAACA
 AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTTCATTCCGAGTGTGTT
 CGGGAGAAGCCACTCCCGATGACATGTCTTGCTGACTATGACAAATGGCTGTGAACCCATCA
 GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA
 GTTCTCCCCAGACCAATGGGGAGGAGTAAGGAGTCTGGATGGGTGAAAAATGAAATATATG
 GTTATTAGGACATATAAAAACTGAAACTGACAACAATGGAAAAAGAAATGATAAGCAAAATC
 CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT
 GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTGGCTGTATCCTTTAT
 CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCAGGCTGCGCACATAGTA
 GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG
 GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAAACACCTCTGTTTTCTTGCTCTATACAG
 CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCATATGGTAGCACAG
 GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTCAAAGAATAAAATCAAATAAAGA
 GCAGGAAAAA

100
 90
 80
 70
 60
 50
 40
 30
 20
 10
 0

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSI ESEDEQK LIEKF IENLLPSDGFWIGLRRREEKQSNSTACQDL
YAWTDGSGISQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPFYMFQWNDRCNMKNNFICKY
SDEKPAVPSREAEGETELTTPVLPEETQEEDAKKTFKESREAA NLAYILIPSIPLLLLLLV
VTTVCVWVICRKRKRQEPDPSTKKQHTIWFSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSES GFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGTCCCTGCTTGTCTGAAGGGCTGGATGTACGC
ATCCGAGGTTCCCGCGGACTTGGGGGCGCCGCTGAGCCCCGGCGCCCGAGAAGACTTGT
GTTTGCTCCTGTCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT
GTTTCAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGGCGGGTGGCCCTTGCCGAGCTGCAGGAGGCCGATGGCCAGTTCGGGTGACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCACCCCCAAACTC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCCTGAAGGGGGCATGTTTGTGCGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTCACCAACCTTCAACCCACAGGAGGTCTTATTCGTTCACATAACATTTTTCGGAATCTG
GAGTCCACCCGTTGTTTGTGCTGGGCTTTTCCAGTGTGAGAAAGAAGGACCCATCATCAT
CCACACTGATGAAGCAGATTGAGAACTCTTGATCCCAACTCCAAAGCTGCTGGAGCCTGA
GGCAGAGAACCAGAGGCCGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTG
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGATACACTGCCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCTGGGGATTTTGGACCACAAATGGCCACCGTTTGTCT
GTTGACCTGACCATGGAACCTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCCAGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAAATAAGTGCCCTTATACAATG

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 52

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVFPQTQFDYTVTNLAGGPKFPYSFYDSQYHETTLKGGMFAGQL
TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFCQCKE
GPIIHTDEADSEVLYPNYQSCWSLRQRTRGRRTASLQPGISEDLKVKVDRMGIDSSDKVD
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFLLHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMS VYTLSPKEYHALCSQTQVMEVGNEE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTCAGCCAGACCTGGATTCTCTAGCGTCTCCATCTGGAGTGC GGCTGGTGGGGGCCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCCTGTGTGATG
ACGGCTGGGACATTAAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAACCAGAGAGCTCTTTCTCCCCA
GTCCAGAGGGGTGT CAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCTGCTCAGGACGAGAAGCAACCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACCTGCAACCATGATGAAGACACGTGGGTCTG
AATGTGAAGATCCCTTTGACTGTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACCTGGGAGAAAAAGGAGGA
CCAGGTGGTATGAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCTGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTG TAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACAGAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTGTTGCCCTGATTCTCAGGCCTTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA
CCACCTTTCTATGTCTCCACATTGCACACAGCAGATCCCAGCCTCCATAATTTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACATA
CACCATTGTCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAAGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAGATAGCAC
TATGTGTTCAAA

FIGURE 54

MALLFSLILAICTRPGFLASFSGVRLVGGLHRCGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEVEYDCSHDEDAGASC
ENPESFSFPVPEGVRLADGPGHCKGRVEVKHQNQWYTVQCQTGWSLRAAKVVCRLGCGRVLC
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG
GDNLCSGRLEVLHKGWVGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,
267-271, 294-298, 316-320, 336-340

N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,
180-186, 263-269, 286-292

Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTTCATCACTCCACAGCCATCCTGCCCTGCTGTTTCGGCTGCCTGGG
CGTCTTCGGCTCTTCGGCTGCTGCAGTGGGTGCGCGGGAAGGCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCCCACTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTCATGGAGACAACTACTTTGGCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCAGCAACCCAGGCTTCTTTGA
CTGCTCGGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAAACTAGAAATAAACATCTCAAAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCACTTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTAC

FIGURE 56

MDFITSTAILPLFLGCLGVFGLFRLLQWVRGKAYLRNAVVTGATSGLGKECAKVFYAAGA
KLVLGCRNGGALEELIRELTASHATKVQTHKPVLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTVDVDRVMETNYFGPVALTKALLPSMIKRRQGHIVAISIQGKMSI
PFRSAYAASKHATQAFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDDTT
AQGRSPVEVAQDVLAAVGKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

FIGURE 57

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
 AAAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAAATTTCTCTGGACATCCTC
 CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTTAA
 GAGGAGAAAATCAGTCACCGGCGAAATCGTGTGATTACAGGAGCTGGGCATGGAATTGGGA
 GACTGACTGCCATATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
 CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGACTGGGTGCCAAGGTTCATACCTTTGT
 GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAGAAGGTGAAGGCAGAAATTG
 GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGTCTACA
 CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
 GGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCCATATGTCACTGTGGCTTCGGCAG
 CTGGACATGTCTCGGTCCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA
 TTTCTAAAAACTTTGACAGATGAACCTGGCTGCCTTACAAATAACTGGAGTCAAAAACACATG
 TCTGTGTCCTAATTTCTGTAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
 CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG
 ATTTTTATTCCATCTTCTATAGCTTTTTTAAACACATTGGAAGGATCCCTTCCTGAGCGTTT
 CCTGGCAGTTTTTAAACGAAAAATCAGTGTAAAGTTTGATGCAGTTATTGGATATAAAATGA
 AAGCGCAATTAAGCACCTAGTTTTCTGAAAACCTGATTTACCAGGTTTAGGTTGATGTCATCTA
 ATAGTGCCAGAAATTTAATGTTTGAACCTTCTGTTTTTTCTAATTATCCCCATTCTTCAATA
 TCATTTTGGAGCTTTGGCAGTCTTCATTTACTACCCTTGTTCCTTTAGCCAAAAGCTGATT
 ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA
 CCAAAATGACTTTATTAAATAAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA
 AATTGTACCATAACCGTTTATTAAACATATATTTTATTTTGTGATTGCACCTTAAATTTTGT
 ATAATTTGTGTTTCTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA
 TGAAGGACTATATCTAGTGGTATTTACAATGAATATCATGAACCTCTAATGGGTAGGTTTC
 ATCCTACCCATTGCCACTCTGTTTCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT
 GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG
 GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAATGGATCACACTTAAAAAAA
 AA
 AA

100
 90
 80
 70
 60
 50
 40
 30
 20
 10
 0

FIGURE 58

MKFLLDIILLPLLVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAACKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 59

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCGCGGCTC
AGGGAGGAGCACCAGCTGCGCCGCACCCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG
TTTCGTGGTCCTGTTGATGCGCTGGCCCCGTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCAGGAGCTGAACATGA
AGAGTTATGCCCGGCTTCTCACCCGTGAATAAGACTTACAACAGCAACCTCTCTTCTGGTTC
TTCCACGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG
AGGTTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTGTCACAAGTAACATGA
CCTTGCCTGACAGAGACTTCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT
AGCACGGGATTTATACAGTGCACTAATTGAGTTTTTCCAGATATTTCCCTGAATATAAAAATA
ATGACTTTTATGTCACTGGGGAGTCTTTATGCAGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCCTCAACCCGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCTGTACCAAAATTGGCT
TGTTGGATGAGAAGCAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
AGGAAGCAGAACTGGTTTGAGGCCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC
AAGTGATCCTTCTTACTTCCAGAAATGTTACAGGATGTAGTAATTACTATAACTTTTTCGGT
GCACGGAACCTGAGGATCAGCTTACTATGTGAAATTTTGTCACTCCAGAGGTGAGACAA
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACCTAGTTGAAAAGTACTTGCAGAG
AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCTGCAGAGCGCTCCTTGATGGGC
ATGGACTGGAAGGATCCAGGAATACAAGAAGGCAGAAAAAAAGTTTGGAAAGATCTTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC
GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA
TTCATTATGGAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCCAAAAGAGAACAT
CAGAGGTTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAA
TTATCTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAAATATCTTGAACAAGTGAGC
TTTTGTTTTTGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
TTATTTAACCTTAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAAGATGTATAAATGA
AATTTTAGGGTCTTGAATAGGAAGTTTAATTTCTTCTAAGAGTAAGTGAAGAAGTGCAAGTTG
TAACAAACAAGCTGTACATCTTTTCTGCCAATAACAGAAAGTTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATTGGATAAGAATAGCTCAATTATCCCAATAAATGGATGAAGCTATAA
TAGTTTTTGGGGAAGATTTCTCAAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA
AAAAATATTATATATAAAGTAAAAAATA

FIGURE 60

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDGQPLFLTPYIEAGKIQKGRELSL
VGPFPLNLMKSYAGFLT VNKTYSNLFFWFPAQIQPEDAPVVWLQGGPGGSSMFGLFVEH
GPYVTSNMTLRDRDFPWT T T L S M L Y I D N P V G T G F S F T D D T H G Y A V N E D D V A R D L Y S A L I Q F
F Q I F P E Y K N N D F Y V T G E S Y A G K Y V P A I A H L I H S L N P V R E V K I N L N G I A I G D G Y S D P E S I I G G
Y A E F L Y Q I G L L D E K Q K Y F Q K Q C H E C I E H I R K Q N W F E A F E I L D K L L D G D L T S D P S Y F Q N V T G
C S N Y Y N F L R C T E P E D Q L Y Y V K F L S L P E V R Q A I H V G N Q T F N D G T I V E K Y L R E D T V Q S V K P W L T
E I M N N Y K V L I Y N Q L D I I V A A A L T E R S L M G M D W K G S Q E Y K K A E K K V W K I F K S D S E V A G Y I R Q
A G D F H Q V I I R G G G H I L P Y D Q P L R A F D M I N R F I Y G K G W D P Y V G

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,
353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

Year	Age	Sex	Weight (kg)	Height (cm)	Body Mass Index (kg/m ²)	Waist Circumference (cm)	Hip Circumference (cm)	Waist-Hip Ratio
1990	20	M	65	175	21.1	85	95	0.89
1991	21	M	68	178	21.5	88	98	0.90
1992	22	M	70	180	21.9	90	100	0.90
1993	23	M	72	182	22.3	92	102	0.90
1994	24	M	75	185	22.7	95	105	0.90
1995	25	M	78	188	23.1	98	108	0.90
1996	26	M	80	190	23.5	100	110	0.91
1997	27	M	82	192	23.9	102	112	0.91
1998	28	M	85	195	24.3	105	115	0.91
1999	29	M	88	198	24.7	108	118	0.91
2000	30	M	90	200	25.0	110	120	0.92

GAGGGGCTTTCCGGCTCCGGAATGGCACATGTGGGAATCCAGTCTTGTGGCTACAACAT
TTTTCCCTTTCTCAAGAGTTCTCAACAGCTGTGTTCTAACAGCTAGTGATCAGGGGTTCTCTT
GCTGGAGAAGAAAGGCTGTGAGGCGACAGCGGGCACTCTCACTCAGGTCAGCCAGCTCCTTG
CCTCTCTGTGGATACAAGACAGATGAGAAAGTGAAGAGATGTCAGCGGAGTGAGGTTGATGGAAG
TCTAAATAGGAAGAAATTTTGTGTGCAATACAGACTCTGGGAGCAGTTGACCTGGAGAGC
CTGGGGGAGGGCCTGCCTAACAGCTTTTCAAAAAACAGAGCGACTTCCACTGGGCTGGAT
AAGACGTGCGCGGTAGGATAGGAAGAGCTGGGTTTAGTCTCAATATCAAAATGACTGGTGCGG
TGAACITCAACAGCTTTTAACTCTCTGGGAGATGAAACAGTAGGCTTAAGGGCCAGAA
TAGAGATGCTTTGTAATAATAAATTTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
GACCAAAATAGATAACAGAGATTCCTTGAACATCTCAAGAGGGAGAAGATATGTTAAAAATA
GAAAAACCAAATGCAGAGAGGAGGACTACAGAGACTAAACAGGATGGGGACCTGGGTCT
AGGCCAGCTGTTTGTCTCCTCCCGGAAATTTTGTGTCTGACCACTCTGCGCTTGGCTTCT
GCGAATCATGTGGAGGCCAACCGGGGAAGTTGGAGCAGCTGAGCACACAGAGGTCGGTTCT
CCTCACGGCCGCCCTCTCAGATGGAAACAGAGGCAGCCTGGCCCCGGGGCCTGGAGGTGG
ACAGCGCTCTGTGGTGCTGCTCTCAGTGGTCTGGGTGCTGTGGCCCCCAGCAGCGCGT
ATGCTCTAGTTTACAGCACTTCCACTCTGAGAATGCTGACTGACCTTCAACCACTTGACCGT
CCACCAAGGACGCGGGGCGCTCTATGTGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA
ACCTGACCATCAGGTTGGCTCTATAAGACAGGGCCAGGAAGGACCAACAGTCTCGTTAACCC
CCCCCTCATGTGCGACGCTCGACGCGAAGTGTCAACCTCAACCAATGTCAACAGCTGCT
CATCATTGACTACTCTGAGAAACCTCTGCTGCGCTGTGGGAGCTCTACCAAGGGGTCTGCG
AGCTGCTGGCGTGGATGACCTGCTCTCTGTTGGAGCCATCCCAAGAAGGAGCACTAC
CTGTCCAGTGTCAACAAGACGGGCAACATGTACGGGGTGATTGTGCGCTCTGAGGTGAGGA
TGGCAAGCTCTTCTCATCGGACGGGTGTGGATGGGAAGCAGGATTACTTCCCGACCTGTCCA
GCGGAAGCTGCCCGGACGACCTGAGTCTCTCAGCTGCTGCATGTAGGATACAGCGAT
TTGTGTCCTCTCTCATCAAGATCCCTTCAGACACCTTGGCCCTGGTCTCCCACTTTGACAT
CTTCTACATCTACGGCTTTGTGATGGGGGCTTGTCTACTTTCTCACTGTCCAGCCGAGA
CCCCTGAGGGTGTGGCCATCAACTCGTGCGGAGACCTTCTTCAACCTCAGCATCTGTGCG
CTCTGCAAGGATGACCCCAAGTTCCACTACATAGTGTCCCTGCGCTTGGGTGACCCGGGC
CGGGTGGAAATACCGCTCTCTGACAGCTGCTTACTGCGCAAGCTGGGAGCTCACTGGCC
AGGCCTTCAATATCACACAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAG
CAGTATCACCCACCGCCGATGACTCTGCGCTGTGTGCTTCCCTATCCGGGCCATCAACTT
GCGATCAAGGAGCGCTGAGTCTGCTGTAACAGGGCAGGGCAACTGGAGCTCAACTGGC
TGCTGGGAAGAGAGCTCCAGTGGCAGGAAGCGCCTGTCCCACTGATGATAACTTCTGTGGA
CTGGACATCAACGACCCCTTGGAGGCTCAACTCGGCTCCAGTGGAGGCGTACCTGTACACCC
CAGCAGGGACCGCATGACCTCTGTGGCTCTCACTGTTTCAACCGCTACAGCGTGGTTTTTG
TGGGGACTAAGAGTGGCAAGCTGAAAAAAGTAAGAGTTATGATGTTAGATGCTCCAAATGCC
ATTCACTCTCTCAGAAAGAGTCCCTCTTGGAGGTAGCTATTGGTGAGATTTAACTATAG
GCAACTTTATTTTCTTGGGAACAAAGGTTGAATGGGGAGGTAAGAAGGGGTAAATTTTGTG
ACTTAGCTTTCTAGCTCTCTCCAGCATCAGTCACTGGGTATGTAAGGAATGCAAGCGTA
TTTCAATATTTCCAAACTTTAAGAAAACCTTTAAGAAGTGATACCTGCAAAAGCAAA

FIGURE 62

MGTLGQASLFAPPGNFYFWSHDHSAFCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLAGCSL
YQGVCCKLLRLDDLFI LVEPSHKKEHYLSSVNKTGMTYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDPRESSAMLDYELHSDFFVSSLIKIPSDTLALVSHFDIFYIYGASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKYHHPPDSSALCAFPRAIRAINLQIKERLQSCYQEGEN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASVYVYNG
YSVVFVGTGSKGLKKVRVYEFRCSSNAIHLLSKESLLEGSYWWRFNRYRLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCGCGCGGCTGAGTGCAGGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCGCATGACCA
 CGTGGAGCCTCCGCGGAGGCGCGGCCGACGCTGGGACTCTGTCTGTGTCTGTCTGTGGCTCTCTGGTGTCTC
 GCAGGCTGGAGTGGAGCACCTGGTCTCTGTGGCTCCGCGCATCGACAGCTGGGGCTCAGGGCCAAAGGCTGGA
 ACTTTCATGCTGGAGGATTCCACCTTCTGGGATCTTGGGGGCTCCATCCACTATTTCGGTGTGCCCGAGGAGTACT
 GGAGGACCGCTCTGTAAGATGAAGGCTGTGGCTTGAACACCTCCACCACTATGTCTCGTGGAACTCTGCATG
 AGCCAGAAAGAGGCAAAATTGACTTCTCTGGGAACCTGGACCTGGAGGCTCTGTCTGTGATGGCCGACAGAGTCTG
 GGCTGTGGGTGATTCTGTGCTCAGGCCCCCTACATCTGCAGTGAGATGGAACCTGGGGGCTTGGCCAGGCTGCTAC
 TCCAGACCTGGATGAGGCTGAGGCACTTACAAGGGCTTCAACCAAGCAGTGGAGCTTTATTTTGACCACT
 TGTATGTCAGGGTGGTGCACCTCCAGTACAAGCGTGGGGGACCTATCATTTGCCGTGAGGTCAGGAATGAATATG
 GTTCTTATATAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTTGTGGAACTGCTG
 TCCGTACTTCAGACAACAGGATGGGCTGAGCAAGGGGATTGTCCAGGGGATTGTGCCCAACTCAACTTGCAGT
 CAACACACGAGCTGCAGCTACTGACCACTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGTGAGGAGT
 ACTGGACGGGGTGGTTGTGACTCGTGGGAGGCCCTCACAATATCTTGGATTCTCTGAGGTTTGAACACCGTGT
 CTGCCATTTGGAGCGCGGCTCTCCATCAACCTCTACATGTTCCACGGAGGACCAACTTTGGCTTCATGAATG
 GAGCCATGCACTTCCATGACTACAAGTCAGATGTCAACGCTATGACTATGATGCTGTGCTGACAGAAGCCGCG
 ATTCACCGCCCAAGTACATGAAGCTTCGAGACTTCTCGGCTCCATCTCAGGCACTCCCTCTCCCTCCCCCACTG
 ACCTTCTCCCAAGATGCGGTATGAGCCCTTAAACGCCAGTCTGTACTGTCTCTGTGGGAGCCCTCAAGTACC
 TGGGGAGGCAATCAAGTCTGAAAAGCCCTCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGACAGTCTCT
 TCGGGTACATTCTCTATGAGACAGCATCACCTCGTCTGGCATCCTCAGTGGCCAGCTGCATGATCGGGGGCAGG
 TGTTTGTGAACACAGTATCATAGGATTCTTGGACTACAAGACAAAGAAATTTGCTCCCTGATCCAGGGTT
 ACACCGTGTGAGGACTCTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCGCGCAAG
 GCTTAATTGGAAATCTCTATCTGAATGATTCAACCCCTGAAAACTCAGAAATCTATAGCTGGGATATGAAGAAG
 GCTTCTTTCAGAGGTTCCGCTGGACAAATGTGNGTTTCCCTCCAGAAACACCCACATATACCTGCTTCTTCTTGG
 GTAGCTTCTCATGCTCCAGCCTTGTGACACCTTCTGAAGCTGGAGGGCTGGGAGAAGGGGCTGTATTCA
 TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCAGAGAAGCGCTTACCTCCACAGTCCCTGGTTGA
 CGAGCGGAATCAACAGGCTCATCGTTTGGAGGACGATGGCGGGCCCTGCATTACAGTTCCAGGAAACCCCC
 ACCTGGGCGAGAACCAAGTACATTAAAGTGAACGGTGGCACCCCTCTCTGCTGGTGGAGTGGGAGTGCCTCGCCCTC
 CTCTGACCTGAAGCCTGTGGCTGCTGCCCCACCCCTCACTGCAAAAGCATCTCTCTAGTAGCAACTCAGG
 ACTGGGGGCTCAGTCTGCCCCCTGTCTCAGCTCAAAACCTTAAGCTCGCAGGGAAGGTTGGGATGGCTCTGGGCC
 TGGCTTTGTGATGATGGCTTCTCAGACCCCTGTCTTGTGCCAGGCTGTGGGCTGTCTCTAGGTTGGGAGC
 AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGTGAAACGTGCCCTTGACCGGAGCGTCAGACCC
 TGCAGCATCTGTGGACTCAGCGTGTCTTTGTCTGGTTCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT
 TTTATCCCGAAATCTGGGTGTGTACCAAGTGTAGAGGGTGGGGAAGGGTGTCTCACCTGAGCTGACTTTGTT
 CTTCCCTTCAACCTTCTGAGCCTTCTTGGGATTCTGGAAGGAACCTCGGCTGAGAAACATGTGACTTCCCTT
 TCCCTTCCCATCTGCTGTCTCCACAGGCTGACAGGCTGGGCTGGAGAAACAGAAATCTCCTCACCTGCGTCTTCC
 CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCTCGGAGAACCATGGCCCATGTCTGCA
 CATCCAGGGAGGAGGACAGAAGGCCAGCTCAGATGTGAGTCTGGCAGAACCATGGCCCATGTCTGCAACATCC
 AGGAGGAGGACAGAAGGCCAGCTCAGATGTGAGTCTGGCAGAACCATGGCCCATGTCTGCAACATCCAGGAGGAGG
 ACAGAAGGCCAGCTCAGTGGCCCCGCTCCCCACCCCCACGCCGACAGCGGGGACAGAGAGCCCTCTCTC
 GAAGTGTGTCAAGTCCGATTGTAGCCTTGTCTGGGGCCAGCCCAACACTTGGCTTGGGCTCACTGTCTGTA
 GTTGAGTAAAGCTATAACCTGGAATCACAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 64

MTTWSLRRRPARTIGLLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIIHYFRVPREYWRDRLLKMKACGLNTLTYYVPWNLHEPERGKFDFSGNLDLEAFVLMA
ABIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTAEVDLYFDHILMSRVVPLQ
YKRGGPILIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKGLSKGIVQGVLAT
INLQSTHELQLLTFFLNFVQGTQPKVMMEYWTGWFDSWGPHNILDSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAHYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGNGQSFGYIILYE
TSITSSGILSGHVHDRGQVFVNTVSI GFLDYKTKIAVPLIQGYTVLRIILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSP LKNFRIYS LDMKKSFQRFGLDKWXS L PETPTLP AFFLGSL SIS
STPCDTFLKLEGWEKG VVFINGQNLGRYWNIGPQKTLYLP GPWLSSGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCGAGGCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAACGCCCATATGGCTCCC
AAGAAGCTGTCTGCTCCCTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCA
GGCAGACACTCGGTCTCGTTCGTAGTGATAGGGGTCATGACCGGTTTCTCCTAGACGGGGCCC
CGTTCGGCTATGTGTCTGCGCAGCCTGCACCTACTTTCGGGTACCGCGGGTGCTTTGGGCGGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACATA
CCACGAGCCACAGCCTGGGGTCTATAAATTAAATGGCAGCCGGGACCTATTGCTCTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTGGTGCATCTAGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGTCTCCCATCTCGTTGCTTCGAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCTTGGCCGAGTGGACTCCTGGTTCAGGCTTTGCTGCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTCAGGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACCTGGCTGGGCTCTTCGCTGCACTGCTAGG
AGAAAAGATCTTGCTCTTCAACACAGATGGGCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATACCCTGTAGATTTTGGCCAGCTGACAACATGACCAAAATCTTTACCTTGCTT
CGGAAGTATGAACCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTGCTGACTGTAACCAAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGGCGATAAGAAGGGACGCTTCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTTGTCTCTTCGAGATGTATCAGCAAGT
TCCAGGAAGTTCTTTGGGACCTTTTACCTCCCCGAGCCCAAGATGATGCTTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAACTCTGCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCATACCATTTTGGAGCCAACCATCTCGGGTGCCAAATATATGGAGTC
CATGACCGTGCCATATGTATGTTGGATGGGGTGTTCAGGGTGTGTTGGAGCGAAATATGAG
AGACAAATATTTTGGAGGGGAACTGGGGTCCAACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCTGTGTAAGCCACCAATTCTG
GGGCAAAACATCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAAGTCTGGATCAATGGGTTTAACTTGGGCGGTCTGGAACAAAGCAGGGGCC
ACAACAGACCCCTCTACGTGCCAAGATTCTGCTGTTTCTAGGGGAGCCCTCAACAAAAATTA
CTATGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTGGATGAAGCCTATC
CTCAATGAGCTAGTACTTTGACACAGGACACATATCAATTCCTTTCAGCTGATACACTGAG
TGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAAGGTAGGCCGGGCATGGTGGCTCATGC
CTGTAATCCCGACCTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTGAGGACTTCAAGA
CCAGCTTGGCCCAACATGGTGAAACCCCGTCTCCACTAAAAATCAAAAAATAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCAGCTACTTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGTCAGTGAGTGAGGTTGTACCACTGCACCTCCAGCCTGGCTGCAGTGA
GACACTCCATCTCAAAAAAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 66

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLCLKMRWSGLNAIQFYVFPWNYHEFPQGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGPLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSVMRHLAFLFRALLGEEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEHPHGPLVNSEYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMVMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPSPKMML
GPVTLHLVGHLLAFLDLCLCPRGPPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVFN
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWYPQAPSGPTFYSKTFPILGSGVGDFTLYL
PGWTKGQVWINGFNLGRYWTQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,
603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGCTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
 ACCCACAATATGCGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
 TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTACAGGATACCTTTGAAGGAATATT
 CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT
 GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTTGGTGTGTT
 CTTGTGAGAAGTTAGTGAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
 AAAAAGTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTCATG
 CTGTCGGGGGTGCCGATGCTGCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC
 AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
 TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
 TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
 CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAAACAATAAGATGATAGGACTTG
 AATCTCTCCGAGAGTTGCGGCACCTTAAGATTTCTCCACGTGAAGAGCAATTTGACCAAAGTT
 CCCTCCAACTTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTACATAATGACGGCAC
 TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTGCTGAGCTGGAACTCCAGA
 ACTGTGAGCTAGAGAGAAATCCACATGCTATTTTACGCTCTCTAATTTACAGGAACTGGAT
 TTAAAGTCCAATAACATTTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAAACGACT
 GACTTGTTTAAATTTATGGCATAACAAAATTTGTTACTATTCTCTCCCTCTATTACCCATGTCA
 AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
 AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
 AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTTGTC
 CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAAGTGCATCACC
 TCACTCCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAAGTGC
 CTTGGACGCGCTGCCAGCCAGCTGGGCCAGTGTGCGGATGTCAAGAAAAGCGGGCTTGTG
 TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAGAGGCATTGAATCAAGACATA
 AATATTCCTTTGCAAAATGGGATTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
 AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
 ATACATCTTTTAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
 GTTCAATGTTTGTAGGGTTTTAAGTCATTCAATTTCCAAATCATTTTTTTTTTTCTTTTGGGG
 AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTCTTTTTTAAATTGTTTGTAACTTGGAT
 GCTGCCGCTACTGGAATGTTTACAATTTGCTTGCCTGCTAAGATAAATGATTAAATTGACATT
 TTCTTACTAAAAA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 68

MAYMLKLLISYISIIICVYGFCIYTLFWLFRIPLKEYSF EKVRESSFS DIPVKNDFAFL
LHMVDQYDQLYSKRFGVFLSEVSENKLR EISLNHEWTF EKLRQHISR NAQDKQELH LFM LSG
VPDAVFDLTDL DVLKLELIPEAKIPAKISQMTNLQELH LCHCPAKVEQTAFS FLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLSEN NKMIGLESRLRELRLKILHVKS NLTKVPSN
ITDVAPHLTKLVIHNDG TKLLVLNLSKKMMNVAE LQNC ELERIPHAIFSLSNLQELDLKS
NNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLES LYFSNNKLES LPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTNLNGQNCITSLP
EKVGQLS QLTQLELKGNC LDR LPAQLGQCRMLKKSGLVVEDHLFD TLPLEVKEALNQDINIP
FANGI

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 326-330

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCCGGCCTTCTCTCGGACCTTGCATTCCATTCCTTTTCATTGACAACTGACTTTTTTATTTCT
 TTTTTTCCATCTCTGGGCGCAGCTTGGGATCCTAGGCGCCCTGGGAGACATTTGTGTTTACACACATAAGGAT
 CTGTGTTTGGGGTTTCTTCTTCTCCCTGACATTGGCATTGCTTAGTGGTGTGTGGGGAGGGAGACCACGTGG
 GCTCAGTGCTGCTTGCCTGACCTATCTGCTTAGGTACATCGAAGCTCTTTTGAGCTCCATACAGTGATATGCTGTCT
 ATCGCTGGTGGTATCTCTGGCGCCCTTGCTCCTGCTGATAGTTGTGCTGCTCTGTCTTTACTTCAAAATACACAAC
 GCGCTAAAAGCTGCAAGAGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAGAGGTGTGGTGGGCGCAAG
 AACAGCCAGGCGCAAAACCAATTGCCACGGAGTCTTGTCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC
 AGTTTGTGATTCCCTGCCACCTTGTCTGTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA
 GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG
 CTCCTCCATCAGTTTCATGAAAAATACTCAGTGCCCTGCTGGGAACACAGCTGCTGGAGATCCCTACAGAGAGCTTC
 CACTGGGGGCAACCCCTCCAGGAAGGAGTTGGGGAGAGAGAACCTCACTGTGGGGAATGCTGATAAACCAAGTCA
 CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGGTGGCTGGAACCTGACGCTTCCCTGGAGGTGTCCAGAA
 GCTGATGTAACACAGAGCCTATAAAAGCTGTGCTGCTTAAGGCTGCCAGCGCCTTGCCAAAATGGAGCTTGTA
 AGAAGGCTCATGCCATTGACCTCTTAATTCTCTCCTGTTGGCGGAGCTGACAAATGGCGGAGGCTGAAGGCAAT
 GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAGCCATGATCCTGCAACTCAATCCC
 AGTGAGAAGCTGCACCTGGACAATAGAAAGCCAGAAAAACAAAGCATCAGAAATATCTTTTCTATGTCCAGCTT
 GATCCAGATGGAAGCTGTGAAAGTAAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCGCTGCTGAGGG
 CAAGTCTGCAGTAAAAACGACTATGTTCTGTATTGTAATCATCATCCAGTACATTGAAGCTTTCAAATAGTTACT
 GACTCAGCAAGAAATCAAAGAACTGTCTTTGTCTTCTACTACTTCTCTCTCTCAACATCTCTATTCCAAATGT
 GGGGGTACCTGGATACCTTGGAGGATCCTTACCAGGCCCAATTACCCTAAAGCGCATCCTGAGCTGGCTTAT
 TGTGTGTGGCACATACAGTGGAGAAAGATTACAAGATAAACTAAACTTCAAAGAGATTTTCTAGAAATAGAC
 AAACCTGTGCAATTTGATTTTCTGCCATCTATGATGGCCCTCCACAGCTCTGGCCTGATGGACAGTCTGT
 GCGCTGAGTCACTCCACCTTTCGAATCGTCATCAAATCTCTGACTGCTGTTGTCTCAGATTAATGCCAATCT
 TACCGGGATTTTCTGCTTCTTCACTCACTAATTTATGCAAGAAACATCAACACTATCTTTAAGCTGCTCTTCT
 GACAGATGAGAGTTATTATAAGCAATCTACCTAGAGGCTTTTAACTCTAATGGGAATAAATTGCACTAAAA
 GACCCAATCTGCAGACCAAAATTAATAAATGTTGGAAATTTCTGCTCCCTTAATGGATGTGGTACAAATCAGA
 AAGGTAGAAGATCAGTCAATTAATCAACCAATATAATCACTTTTCTGCTATCCTCAACTTCTGAAAGTATCAAC
 CGTCAGAAAACACTCCAGATTTATGTGAAGTGTGAAATGGGACATAATTTACAGTGGAGATAATATACATAACA
 GAAGATGATGTAATAAAGTCAAAATGCACTGGGCAATATAACACCGCATGGCTCTTTTGAATCCAAATCA
 TTTGAAAGACTATACTTGAATCAACATATTAATGTGGATTTGAACCAAACTCTTTTGTCAAGTTAGTCTGCAC
 ACCTCAGATCCAAATTTGTGTGTGTTTCTTGATACCTGTAGAGCTCTCCCACTCTGACTTTGCATCTCCAACT
 TACGACCTAATCAAGATGGATGTAGTCGAGATGAAACTTGAAGGTGATCCCTTATTGGACACTATGGGAGA
 TTCCAGTTTAATGCTTTAAATTTCTTGAGAAATGATGAGCTCTGTATCTGCAGTGTAAGTTTGTGATATGTGAT
 AGCAGTGACCAACAGTCTGCTGCAATCAAGTTGTGTCTCCAGAGCAACAGAGACATTTCTTCATATAAATGG
 AAAACAGATTCATCATAGGACCACTTCGCTGAAAAAGGATCGAAGTGCAAGTGGCAATTGAGATTTGAGCAT
 GAAACACATGCGGAGAAACCTCCAAACAGCCCTTTCAACAGTGTGCATCTGTTTCTTCTCATGGTCTAGCTCTG
 AATGTGGTGATCTAGCGCAATCAGAGTGGGCAATTTGTAATCAACGGGCAGACTACAAATACAGAAAGCTG
 CAGAACTATTAATAACAGGTCCAACCTTAAGTGAGACATGTTTCTCCAGATGCCAAAGGAATGCTACCTCGT
 GGCTACACATTTATGAATAAATGAGGAAGGCGCTGAAAGTGACACAGGCTGATGTAATAAA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 70

MELVRRMLPLTLLILSLCLAEHLTMAEAGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI
ERPENKSIRIIIFSYYQLDDPGSCSENIKVFDGTSSNGPLLGGVCSKNDYVPVFESSSSTLT
FQIVTDSARIQRTVFVFYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPPELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSNSLTVVLS
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVVIISKSYLEAFNSNGNNLQKDPCTCRP
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST
VEIIYITEDDDVIQSONALGKYNTSMALFESNSFEKITLESPYYVDLNLQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDFASTYDLIKSGCSRDETCVYPLFGHYGRFQFNAFKFLRSMSSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDISIIGPIRLKRDRSASGNSGFQHETHA
BETPNQPFNSVHLFSFMVLALNVVTATITVRHFVNQRADYKYQKLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
 GGACATGCGGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGCTGCTGC
 TGTGCGCGCCGCGCGCTGCCCTGCCACAGCGGCCACGCGCTTCGACCCCATCTGGGAGTCC
 CTGGACGCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTTCATCCACTG
 GGGAGTGTTTTCCGTGCCAGCTTCGGTAGCGAGTGGTTCTGTTGGTATTGGCAAAAGGAAA
 AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCTCCTAGTTTCAAATATGAAGAT
 TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAAGTGGGCAGATATTTTTTCAGGC
 CTCTGGTGCCAAATACATTGTCTTAACTTCCAACATCATGAAGGCTTTACCTTGTGGGGGT
 CAGAATATTCGTGGAACCTGGAATGCCATAGATGAGGGGCCAAAGAGGGACATTGTCAAGGAA
 CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGACTGTACTATTCCCTTTTTGA
 ATGTTTTCATCCGCTCTTCTTGAGGATGAATCCAGTTTCATCCATAAGCGGCAATTTCCAG
 TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACATCAGCCTGAGGTTCTGTGG
 TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
 ATATAATGAAAGCCAGTTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
 TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA
 CATAAATGGGAAAACCTGCATGACAATAGACAACTGTCTGGGGCTATAGGAGGGAAGCTGG
 AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCACTTGTAGAGACAGTTTCATGTG
 GAGGAAATCTTTTGATGAATATTGGGCCACACTAGATGGCACCATTCTGTAGTTTTTGAG
 GAGCGACTGAGGCAAGTGGGCTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
 TACCTGGCGATCCCAAGATGACACTGTACCCCGATGTGTGGTACACATCCAAGCCTAAAG
 AAAAAATTAGTCTATGCCATTTTTCTTAAATGGCCACATCAGGACAGCTGTTCCTTGGCCAT
 CCCAAAGCTATTCTGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG
 GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC
 CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
 ATGCTGCAAGTTATGTCTAAGGCTAGGAACATCAGGTGTCTATAATTGTAGCACATGGAGA
 AAGCAATGTAACTGGATAAGAAAATTTTGGCAGTTTCAGCCCTTTCCCTTTTTCCCACTA
 AATTTTTCTTAAATACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC
 TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAAATTTTTCACATTATAGTAG
 CAAGGAATTTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCATG
 ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCAATGCTAGTCAAT
 TTTTTTTTGTGCCAACATCATAGAGTGATTTACAAAATCCTAGATGGCATAGCCTACTACA
 CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
 AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAAACATATGGAAAACATAGAGAAG
 GTACAGTAAAAAATCTGTAATAAATAGGTGCACCTGTATAGGCGACTTACCACGAATGGAG
 CTACAGGACTGGAAGTTGCTCGGGTGAGTCAGTGAGTGAATGTGAAGCCCTAGGACATTA
 TTGAACACTGCCAGAGCTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA
 GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT
 TTTAAACCTTTTTGGCTCTTTGTAATAACACTTAGCTTAAACATAAACTCATTGTGCAA
 ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLLPPPPCPAHSATRFDPWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYQKEKIPKYVEFMKDNYPSPFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRDRLRFGLYSLFEW
PHPLFLEDESSSFHKRQFPVSKTLPPELYELVNNYQPEVLWSDGDDGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGNNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVPDVMYTSKPKKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKKGWALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,
375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT
TTCCATCCAGGTGTCTATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCCAGC
AGCTGAATTTTCACAGAAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA
TGGATTCTGGGTCTCTCTAGGATTAGCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTGGGAAGGTTCCAGTGAAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGAT
ACTTGGACTAATCTGTCGATTCAGAAAATTATCACCACCAAGATCCCATATTCAACACTCA
AACTGCAACACAAAACAGAAATTTATTGTCACTGACAGTACCTACTCAGTGGCATCCCTT
ACTCTACAATACCTGCCCTACTACTACTCTCTCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAATTGATTTGTGTACAGAAATTTTATGAAACTAGCACCATTGTCTACAGAAAC
TGAACCATTTGTTGAAAAAAGAGCAGCATTCAGAAATGAAGCTGCTGGGTTTGGAGGTGTC
CCACGGCTCTGCTAGTGTCTCTCTCTCTCTTGGTGTGTCAGCTGGTCTTGGATTTTGC
TATGTCAAAGGTATGTGAAGGCCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAATGAT
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAATGAGGAATCAAAGA
AACTGATAAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGCATGCTCGGAA
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACCTGAGGCTGGTTCTTTTCATGCTCC
TTACCTTGCCCCAGCTGGGGAAATCAAAGGGGCCAAAGAACCAAGAAAGTCCACCCCTT
GGTTCCTAAGTGAATCAGTCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAATGC
CCTTCTCCTTATTGTAAACCTGTCTGGATCCTATCCTCTACCTCCAAAGCTTCCCAAGGCC
TTTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTGTCAAAGTGCAA
GGACCTAAAACTCTCATCAGTATCCAGTGGTAAAAAGGCCCTCTGGCTGTCTGAGGCTAGG
TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGAGCTCAGAC
CCTTTCTTCACTCTGAAAGAGAAACACGTATCCCACTGACATGTCCCTCTGAGCCCGGTA
AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCGAAAGCCATGGAGATTCTCATAAATTGAG
ACCTAATCTCTGTAAAGCTAAAATAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT
GTCAGCAGGAGCTGTAAACACAGACAGGGTCAAAGTGTCTCTGAAACATTTGAGTTGGA
ATCACTGTTTAGAACACACACTTACTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAACAAAAATAAAACTCTTATAAATTTCTATTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATCTCTCAAATAATGACATAGTAGAACGCTATCTGGGAAGCTATTTTCTCA
GTTTGTATTTCTAGCTTATCTACTTCCAACTAATTTTATTTTGTGCTGAGACTAATCTT
ATTCTATTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG
TACATTGTTACCTCTATATACCAAGCACATTTTAAAAGTGCCATTAAACAAATGTATCACTA
GCCCTCCTTTTCCAAACAAGAGGACTGAGAGATGCAGAAATATTGTGACAAAAAATTAA
AGCATTAGAAACTT

FIGURE 74

MARCFSLVLLLTISIWTRLLVQGSRLABELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAFPFTNKNQKEMIETKVVKEEKANDSNPNEESKKTDKNPESKSP
SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTGGGTGCCGCGACTTTCACGATGG
 CTCGCCCAACCTTACTACCTTCTGTGCGCCCTGCTCTCTGCTGCCTTCTACTCTGTGAGGAA
 ACTGCCGCGCTCTGCCACGGTCTGCCCACCAACGCGAAGACGGTAACCCGTGTGACTTTG
 ACTGGAGAGAAGTGGAGATCTGTATGTTTTCTCAGTGCCATTGTGTATGATGAAGAACCGCAGA
 TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT
 TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC
 TGATGACGTGCAAAACCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA
 ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGAATTGTGGAGTTCTTTGCCAA
 TTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAAC
 GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGTATACTGATGTTAGTACGCGGTAC
 AAAGTGAGCACATACCCCTCACCAAGCAACTCCCTACCTGATCCTGTTCCAAGGTGGCAA
 GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG
 AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAAACTATCA
 AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTGTGAGA
 TGGGGAAAACAAGAAGGATAAAATTAAGATCCTCACTTTGGCAGTGC'TTCCTCTCCTGTCAATT
 CCAGGCTCTTTCATAACCACAAGCCTGAGGCTGCAGCCTTNNATNNATGTTTTCCCTTTGG
 CTGNGACTGNGTGGGGCAGCATGCAGCTTCTGATTTTTAAAGAGGCATCTAGGGAATTGTCAG
 GCACCTACAGGAAGGCCTGCCATGCTGTGGCCAAGTGTTCCTGAGGCAAGAAAGAGATC
 TCATAGGACGGAGGGGAAATGGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTGCTTATC
 AGCTATTGAGACATCTCCATGGTTTCTCCATGAACTCTGTGGTTTCATCATTCCTTCTTAG
 TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCTAAGGTAAGATGCTGGGGTATAGAA
 CGCTAAGAATTTTCCCCAAGGACTTTGCTTCCCTAAGCCCTTCTGGCTTCGTTTATGGTC
 TTCATTAAGATATAAGCCTAACTTTGTGCTAGTCTCAAGGAGAAACCTTTAACCACAAAG
 TTTTTATCATTTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
 TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCTGGACTTTCAC
 TAAACCTCTGCATACCTCCCAACCCAGTTGATGGCTTTCGTAATAAAAAGATTGGGATT
 TCCTTTTG

FIGURE 76

MAVLAPLIALVYSVPRLSRWLAQPPYYLLSALLSAAFLIVRKLPPLCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYN
TGLNFGKVDVGRYTDVSTRYKVSTSP LTKQLPTLILFQGGKEAMRRFQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

FIGURE 78

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD
KEHDVSLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR
HRRLEHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 80

MMWRPSVLLLLLLLRHGAQ GKPS PDAGPHGQGRVHQ AAPLS DAPHDDAHGNFQYDHEAFLGR
EVAKEFDQLTP EESQARLGRI VDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT
YDTRDRGRVGEELRNATYGHYAPGEEFHDVEDAET YKKMLARDERRFRVADQDGD SMATRE
ELTAFLHP EEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEP AWVQTERQQ
FRDFRDLNKDGHLDGSEVGHVWLPPAQDQPLVEANHLLHESD TDKGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,
291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

FIGURE 81

GGGGCCTTGCCCTTCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG
GCGGCGGGGCGCGGGTGCAGGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTGCGTCCAG
CTCTGTCTGTCTGTTTGGCGCCCCCGCTCCCGCGGTGCAACCGATCCTG
GGCTTCGCTCGATTTCGCGCGAGGCGCCTCCAGACCTAGAGGGGCGCTGGCTGGAGCAG
CGGGTCGTCTGTCTCTCTCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGC GGGGCTCT
GAGGAGGTGACGCGCGGGGCTCCCGCACCTTGGCCTTGCCCGCATCTCCCTCTCTCCAG
GTGTGAGCAGCCTATCAGTCACTATGTCGCGAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG
TGTCTGCTGTCTGCTGCCGGGGCGCGGGCAGCGAGGGAGCGGCTCCATTGTCTATCACATG
TTTTACCAAGGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCTCTGCCAGGGGGCTGCC
CTCTTGAGGAATTTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC
TGGTTCGAGAAAACATTCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTCTTAGAT
GGTCTGCTTCTTTACAGTAACATAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
GTGTCCACAGCACATCCACCAAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACCTGG
CAATAAAGATTGTAAGCAGACATTGCATTCTGATTGATGGAAGCTTTAATATTGGGCAGC
GCCGATTTAATTTACAGAAGAATTTTGTGGAAAAGTGGCTCTAATGTTGGGAATTGGAAACA
GAAGGACCACATGTGGGCTTGTTCAGCCAGTGAACATCCAAAAATAGAATTTTACTTGAA
AACTTTACATCAGCCAAAGATGTTTTGTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA
ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTTCTCACGGTAGATGCTGGA
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATGATGGTTGGCCTTCTGATGACAT
CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA
AGCCTATCCCTGAAGAAGTGGGGATGGTTCAGGATGTCACTTTGTTGACAAGGCTGTCTGT
CGGAATAATGGCTTCTTCTCTTACCACATGCCCACTGGTTTGGCACCACAAAATACGTAAA
GCCTCTGGTACAGAAGTGTGCACCTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
CAGTGAACATTGCCCTTTCTAATTGATGGCTCCAGCAGTGTGGAGATAGCAATTTCCGCGCTC
ATGCTTGAATTTGTTTCCAAACATAGCCAAGACTTTGAAATCTCGGACATTTGGTGCCAAGAT
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCACTTTCAGTGACTATAGCACC
AAGAGAATGCTCTAGTGTCTATCAGAAAACATCCGCTATATGATGGTGGTGAACAGCTATGGT
GATGCCATTTCTCTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAGAA
CTTCTCTAGTAATTGTACAGATGGGCAGTCTCATGATGATGTCCAAGGCCCTGCAGCTGCTG
CACATGATCGAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCAAGAGAGTTCACAGGATT
AGAACCAATTGTTTCTGATGTCTATCAGAGGCATTGTAGAGATTTCTTAGAATCCAGCAAT
AATGGTAAACATTTTGACAACTGAAAGAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT
CTCATAAATACTGAAATGCTTTTAGCATACTAGAATCAGATACAAAACATTAAGTATGTCAAC
AGCCATTTAGGCAAAATAAGCACTCTTTAAAGCGCTGCCTTCTGGTTACAATTTACAGTGT
ACTTTTGTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA
GATAATGTGGATTAAACCTTAAAGAGTTCTAACCATGCTCAATAGTACAGATATGCAAA
TTCCATAGCTCAATAAAGAATCTGATACTTAGACCAAAAAAAAAA

FIGURE 82

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFATKEVGFRRGNSNTGKAL
KHTAQKFFTVDAGVRKGI PKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVT FVDKAVCRNNGFFSYHMPNWF GTTKYVKPLVQKLC THEQMMCSKTCYNSVNI AFLI
DGSSSVGDSNFRMLLEFVSNI AKT FEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAHADAGITI
FSVGVAWAPLDDLDKMASKPKESHAF FTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,
425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCGCCACCGCGCGCTCCCGCATCTGCACCCGAGCCCC
GGCGGCTCCCGGGCGGAGCGAGCAGATCCAGTCCGGCCCCGAGCGCAACTCGGTCCAGTCG
GGCGCGGCTCGCGGCGCAGAGCGGAGATGAGTCAGCGGCTTGGGGCCACCTGCTGTGCTGC
TGCTGGCGGCGGCTCCCGACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCCGGCCGGCTCTCAGTACCCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCGA
GGTTGAGGAATGATGGAGGACACGCAGCACAAATTGCGCAGCGGCTGGAAAGAGATGGAGG
CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTTGCAAACTTACCTCCAGCTAT
CACATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACACAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCATCATCTG
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCAGC
ATGTACTGCCAGTTTGCAGCTTCCAGTACACCTGCCAGCCATGCCGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAAGTGTCTGTGGAGACAGCTGTGTGTCTGGGGTCACTGCACCAAAA
TGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACAACAGAGGGACTGCCAGCCGGGGCTG
TGCTGTGCCCTCCAGAGAGGCTGTGTTCCCTGTGTGCACACCCCTGCCGTGGAGGGCGA
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACTGGGAGCTAGAGCCTGATG
GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCTCCTCTGCCAGCCCCACAGCCACAGCTTG
GTGTATGTGTGCAAGCGGACCTTCTGTGGGAGCCGTGACCAAGATGGGGAGATCCTGTGCC
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCTGAGGAGGAGTGCAGCAGGAGCTGG
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGT
GCACTGCTGGGAGGGGAAGATTTAGATCTGGACAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTAGGCGTGGGCTGACAGGCTTCTTCTTACA
TCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCAATTTGTTTCAGC
TCCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC
TGCAGGAGCAGTTTGCACCCCTGTCCAGATTATTTGGCTGCTTTGCCCTTACCAGTTGGCAG
ACAGCCGTTTGTCTACATGAGCTTTGATAATTGTTGAGGGGAGGAGATGGAACAATGTGG
AGTCTCCCTCTGATTGGTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAA
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTCTTTCCATGGGCATAGGTAAGCTG
TGCCCTTCAGCTGTGTCAGATGAAATGTTCTGTTCAACCTGCATTACATGTGTTTATTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCCTGGGGAGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACAGCCTTGGTGCCACCAA
AAGTGTCTCCCCAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG
GTCAAACTAATTTCTCACATCCCCTCTAAAGTAACACTACTGTTAGGAACAGCAGTGTTCAC
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACATGTCTCCTCTTTGGCAGT
TGCAATAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA
GTACTTAGGTAAATTTAGGGCGAGGATTAATAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAGACAATTATCAACCACGTGGAGAAAAATCAACCGAGCAGGGCTGTGTGAAACATGGTT
GTAATATGCGACTGCGAACACTGAACCTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCA
TGGACTGTTGCCACCATGTATTATCCAGAGTCTTTAAAGTTTAAAGTTGCACATGATTGTA
TAAGCATGCTTTCTTGTGTTTAAATTTATGATAAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 84

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVVEELMEDTQ
HKLRSAVEEMEAEEAAAKASSEVNLANLPSPYHNENTNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFYTCQPCRQGRMLCTRDECCG
DQLCVWGHCTKMATRGSGNTICDNQRDCQPLCCAFQRLGLFPVCTPLPVEGELCHDFASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEET

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACTCTCTGGGCTCAGAAGGACTCTG
AAGATAACATAAATTTCAGCCCATCCACTCTCCTTCCTCCCAAACACACATGTGCATGTACACACACATACAC
CACACATACACCTTCTCTCTCTTCACTGAAGACTCACAGTCACCTCACTCTGTGAGCAGCTTATAGAAAAGGACAC
TAAAGGCTTTAAGGACAGGCTGGCCATTACCTCTCGAGCTCCTTTGGCTTTGTGAGTCAAAAACATGGGAGGGG
CCAGGCAAGGTGACTCACACTCTGTAATCCAGCATTTTGGGAGACCGAGGTTGAGCAGATCACTTGAGGTCAGGAG
TTCGAGACAGCCTGGCCACATGGAGAAACCCCATCTCTACTAAAAATAAAAAATTAGCCAGGAGTGTGTGGC
AGGTGGCTGTAAATCCAGCTACTCAAGTGGCTGAGCCAGGAGAAATCGCTTGAATCCAGGAGGCGGAGGATGCACT
CAGCTGAGTGACCCGCTGCACCTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAAACAAACAAACACGGGAGGA
GGGTAGATACTGCTTCTCTGCAACCTCTTAACTCTGCATCTCTTCTCCAGGCTGCCCTTGATGGGGCTTG
GCAATCACTGAGCAGGCCAGCCCCAGAGGACAAGGAAGAGAAGGCAATATTGAGGAGGGCAAGAGTGAACGCCCG
GTGTAGAAATGACTGCCCTGGGAGGTTGGTCTTGGGCCCTGGCAGGTTCTGACCTTTACCTTGCAAAACACA
AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCTGCCCTGCAGCTCCACATAGAGGCTTCTCGTGGCCCC
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCTTGGCATGTTCCCTGCCCCCTCA
GTGTGCTGCCAGATCCGGCCCTGGTATACGCCCCGCTGCTCTACCGCGAGGCTACCACTGTGGACTGCAATGA
CCTATTCTGACGGCAGTCCCCCGGCACTCCCGCAGGGACAAGACCCCTGCTCTGCAGAGCAACAGCATTTGT
CCGTGTGGACAGAGTGAAGTGGCTACCTGGCCAACTCTACAGAGCTGGACCTGTCCAGAGAACAGCTTTTCGGA
TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGTCTGAGCCTGCACCTAGGAGGAGAACAGCTGACCCGGCT
GGAGGACCAAGCTTTCTGAGGCTGGCCAGCTTCCAGGAATCTATCTCAACCAACAAAGCTTACCCGCTACGC
CCCCAGGCGCTTTTCTGGCCCTCAGCAACTGTCTGGGCTGCACCTCAACTCAACTCTCTGAGGCACTGACAG
CCGCTGGTTTGAATGCTGCCCACTTGGAGATCTCATGATTGGCGGCAACAGGTAGATGCCATCTGGACAT
GAACCTTCGGCCCTGGCCAACTGGCTGAGCTGGTGTGCTAGCAGGCAAGAACCTGCGGGAGATCTTCGACATATGC
CCTGAGGGGGCTCAAGAGCTGGAGAGCCTTCTCTTCTATGACAACAGCTGGCCCGGGTCCGACAGGCGGGCACT
GGAGAGGTGCCCGGCTCAAGTCTTAGACTCAACAGAACCCGCTCCAGCGGTTAGGGCGTGGGGAGCTTTGC
CAACATGCTGCACCTTGAAGAGCTGGGACTGAACAAACATGAGGAGAGCTGGTCTCCATGCACAACTTTGCCCTGGT
GAACCTCCCCAGCTTAAAGAGCTGGACATCAACAAATAACCCACGGCTCTCTTCATCAACCCGCTGGCTTCA
CCACTTGGCCAGATGGAGACCTCATGCTCAACAAACAGCTCTCAGTGCTTGCACAGCAGACGGTGGAGTC
CCTGCCCAACTCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTCTATCCGCTGGCCATGC
CAGGGCAACCGGTGTCCGTTCTATCGAGCCGCAATCCACCTGTGTGCGAGGCTCCGGACCTCCAGCGCTCCCC
GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGCACTGTTTGGCCCTCATCTCCCAAGAGCTTCCCCCAAG
CCTCCAGGTAGCAGTGAGAGAGGTGGTGTGCAATGCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG
GGTCACTCCAGCTGGGCTTCACTGACACCTGCCATGCAGGCAAGAGGTACCGGGTGTACCCCGAGGGGACCT
GGAGCTGCGGAGGTTGACAGCAGAAGAGGCGAGGCTATACACTGTGTGGCCAGAACTTGGTGGGGCTGACAC
TAAGACGGTTAGTGTGGTTGTGGGCGGTCTCTCTCCAGCAGGCAAGGACAGAGGACAGGGGCTGGAGCTCG
GGTGAGGAGACCCACCCCTATCACATCTGTCTATCTGGGTACCCCAACACAGTGTCCACCAACCTCAC
CTGGTCCAGTGCTCTCTCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCGCTGCTCTGGGGAACCCACAGCTA
CAACATTACCCGCTCTTCAAGGCCAGGAGTCTGGGCTGCTCTGCAAGTGGCTTTGCTGATGCCCAACCCA
GTTGGCTTGTATGGGCCAGGACCAAGAGGCCACTTCTTGCCACAGAGCTTAGGGGATGCTCCTGGGCTCAT
TGCCATCCTGGCTCTCGCTGTCTTCTCTGGGCACTGGGTAGCGGGCCACTTGGCAAGGCCAACCCAGGAA
GGGTGTGGTGGGAGGCGGCTCTCCTCCAGCCTGGGCTTCTGGGCTGGAGTGCCCTTCTGTCCGGGTTGT
GTCCTTCCCTCTGCTCCGCTGGAAATCCAGGAGGAAGCTGCCAGATCTCAGAAGGGGAGACATGTTGCC
ACCATTTGCTCAAAATCTTGAAGCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTACCAA
AGAGAAGCAGTCTGGCCAGATGCCCTGCCAGGAAGGGACATGAGACCGCTTGGAGGCTGGAGCTGGGC
CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGA AAAAGTTGCCCTTACCTCTAGGGTCA
CCTCTGCTGCATCTCAGGAACATCTCAAGGAACAGGAGGACTTTGGCTAGAGCTCTCTGCTCCCATCTT
CTCTCTGCGAGAGGCTCTGGGCTGGCTTGGCTGCTCCCTACCTGTGTCCCGGGCTGCACCCCTTCTCTCTC
TCTTCTCTGTACAGTCTGATGCTTGTCTGTGCTCTGGGCAAGGGCTGAAGAGGCACTTCATCTCAC
CTGGGGGGCTGCCCTCAATGTGGGAGTGACCCAGCCAGATCTGAAGGACATTGGGAGGGATGCCAGGAA
CGCTCATCTCAGCAGCTGGGCTCCGATTCAGAGCTGACTTCTATAGGCAATTTGTACTTGTGGAGAA
ATGTGTCACTCCCCCAACCCGATTCACCTCTTCTCTCTGTTTGTAAAAATAAAAAATAATAACATAAA
AAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 86

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA
VPPALPAGTQTLLQLQSNISIVRVDQSELGYLANLTELDLSQNSFSRDARDCDFHALPQLLSLHL
EENQLTRLEDHSPAGLASLQBELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE
MLPNLEILMIGGNKVDAIILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLSELSFYDNO
LARVPRRALEQVPGLKFLDLNKNPIQRVGPGDFANMLHLKELGLNMEELVSIDKFALVNL
ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNALSALHQQTVESLPNLQEVGLHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLFPVREVPFREMTDHLPLISPRSFPPSLQ
VASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAEEAGLYT
CVAQNVLVGADTKTVSVVVGALLQPGRDEGGGLELRVQETHPHYHILLSWVTPPNTVSTNLTW
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLLAAGLAHLGTGQPRKGVGGRRPLPPAWAFWGSAPSVRV
SAPLVLWPNPGRKLPRSSGEETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,
243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

[illegible]

FIGURE 88

MRQTIKVIKFILIIICYTVYYVHNKFDVDCITVDIESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRQLNLNNEWTLDDKLRQLTKNAQDKLELHLFMLSIGIPDTVFDLVELEV
LKLELIPDVTIPPISIAQLTGLKELWLYHTAAKIEAPALAFIRENLRALHIKFTDIKEIPLWI
YSLKTLBELHLTGNLSAENNRYIIVDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIIISFQ
HLHRLTCLKLWNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKRLRYLDLSHNNLT
FLPADIGLLQNLQNLAITANRIETLPPELFPQCRKLRLHLGNNVLQSLPSRVGELTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCCGT
 CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCTGAACGCAGGAGCTGTCAATTGACT
 GGCCCAACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG
 TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCCTGGTCAT
 GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTGGAAAATTTGAGGAAATTGGGC
 CCCTTGACAGTGATCTCAAACACGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT
 GTGGATAATCCCGTGGGCATGGGTTCAAGTTATGTGAATGAGTAGTGGTGCTATGCCAAGGA
 CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG
 AATTCCAGACAGTTCATTCTACATTTTCTCAGAGTCTATGGAGGAAAAATGGCAGCTGGC
 ATTGCTCTAGAGCTTTATAAGGCCATTAGCGAGGGACCATTCAAGTGCAACTTTGCGGGGGT
 TGCCCTTGGGTGATTCTCGGATCTCCCTGTTGATTGGTGTCTCCTGGGGACCTTACCTGT
 ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA
 CTGAATGCCGTAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
 GATCATTGAACAGAACACAGATGGGGTGAACCTCTATAACATCTTAACATAAAGCACTCCCA
 CGTCTACAATGGAGTCGAGTCTAGAATTACACAGAGCCACCTAGTTTGTCTTTGTGACGCGC
 CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA
 GCTCAAAATTATTCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGCTCTTTGTGAACA
 TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGTGAGGAGCAGGGATC
 AACGTGACGGTGATAATGGACAGCTGGATCTCATCGTAGATACCATTGGGTGAGGAGGCCTG
 GGTGCGGAAACTGAAGTGGCCAGAACTGCCATAATTAGTCAGCTGAAGTGAAGGCCCTGT
 ACAGTGACCCTAAATCTTTGGAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTT
 TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT
 GATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
 TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
 AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTTGTCTCT
 GGAGGCAATTGGAATATTCTGCTTCTTAAAAAACCTAAGATTTTTTAAAAAATTGAT
 TTGTTTTGATCAAAATAAAGGATGATAATAGATATTA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 90

MELALRRSPVPRWLLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSLKPRKTTWLQAASLLFVDNPGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFPSCHKEFQTVPFYIFSESYGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNGLYRE
ATELWGKAEMII EQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDAL
QLMNGPIRKKLKIIPEDQSWGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP
SDQGDMAKKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,
187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 92

MGARGALLLALLLARAGLRKPESQEAAPLSGPGRRVITSRIVGGEDABLGRWPWQGSRLRW
DSHVCVSVLLSHRWALTAHCFETYSDLSDPGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDI FGDMVCAGNAQGGKDACFGDSGGPLACNKNG
LWYQIGVVSWSWVGCGRPNRPGVYTNISHHFEWIKLMAQSGMSQPDPSWPLLFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,
259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLNVADLVRPSPLTLHTVQKWL LAAGA QKCHSVITQDFLTCWL
SIRQAE LLLPGA E FHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGG LHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQY LMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPSSSYFNASGRAYPDVAALSDGYWVVSNRVPIPVWSGTSASTFVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSPGWD FVTGWGTPTS QLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,
521-527, 533-539, 549-555

FIGURE 95

GCCGCGGCTCTCTCCCGCGCCACACCTGTCTGAGCGGCGCAGCGAGCCGCGGCCCGGC
 GGGTGCTCGGCGCGGAACAGTGCTCGGCCATGGCAGGGATTCCAGGGCTCCTCTTCTCTC
 TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCCCTGGAACCCACTTG
 GCCTGCATACCGCTCCCTGTCGTCTTGCCCCAGTCTACCTCAATTTAGCCAAGCCAGACT
 TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCAGTGTCTAAGGGAAC
 CCATGCCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG
 CAGCCGCACAGAGACGCAGGTGGGCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAAC
 ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGC
 AGGTTTCAGCATTTTTGGGAAGGACTTCTGTCTCAACTACCTTTTCTCAACATCAGTGAAGTT
 ATCCAGGGCTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCTCAGCTGCCCACTGCA
 TACAGATGGAAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC
 AAGTTTAAAGATGGTGGTTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA
 ATTTTCAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA
 ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCACAAGAGAAAA
 TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAAATTCATTCTC
 TGGTTATGACAATGACCGACCAGGCAATTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA
 CCTATGACTTGCTCTACCAGCAATGCGATGCCAGCCAGGGGCCAGCGGGTCTGGGGTCTAT
 GTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGGAGCGAAAAATTTATTGGCATTTTTTCAGG
 GCACCACTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTGAGATCACTC
 CTCTCAAAATATGCCAGATTTGCTATTGGATTAAAGGAACTACCTGGATTGTAGGGAGGGG
TGACCACAGTGTTCCCTCCTGGCAGCAATTAAGGCTCTTCATGTTCTTATTTTAGGAGAGGCC
 AAATTTGTTTTTTGTTCATTTGGCGTGACACGTGTGTGTGTGTGTGTGTGTGTGAAGTGT
 CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAACTG
 GTTTGTGTATCATATCATATATCATTTAAGCAGTTTGAAGGCATACTTTTGCATAGAAATAA
 AAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG
 CAAACTTTGATTTTTATTTCATCTGAACCTTGTTTCAAAGATTTATATTAAATATTTGGCATA
 CAAGAGATATGAAAAAAAAAAAAAAAA

100
 200
 300
 400
 500
 600
 700
 800
 900
 1000
 1100
 1200
 1300
 1400
 1500
 1600
 1700
 1800
 1900
 2000
 2100
 2200
 2300
 2400
 2500
 2600
 2700
 2800
 2900
 3000
 3100
 3200
 3300
 3400
 3500
 3600
 3700
 3800
 3900
 4000
 4100
 4200
 4300
 4400
 4500
 4600
 4700
 4800
 4900
 5000
 5100
 5200
 5300
 5400
 5500
 5600
 5700
 5800
 5900
 6000
 6100
 6200
 6300
 6400
 6500
 6600
 6700
 6800
 6900
 7000
 7100
 7200
 7300
 7400
 7500
 7600
 7700
 7800
 7900
 8000
 8100
 8200
 8300
 8400
 8500
 8600
 8700
 8800
 8900
 9000
 9100
 9200
 9300
 9400
 9500
 9600
 9700
 9800
 9900
 10000

FIGURE 96

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS
SSCGPQCHKGTPLPTYYEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKS
RRKRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAACHIDGKTYVKG
TQKLRVGFLLKPKFDGGRGANDSTSAMPEQMKPQWIRVKRTHVPKGWIKGNANDIGMDYDYA
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD
AQPGASGSGVYVRMWRKQKQWERKIIGIFSGHQWVDMNGSPQDFNAVVRITPLKYAICYW
IKGNYLDCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

FIGURE 97

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCTGCTCCCCGCCCCACCAGCCATG**GT**GGTTT
 CTGGAGCGCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
 GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA
 GCAGCTGAACCGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
 GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
 ACTGCTGCCACTGTTTCAAGGACAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTGGG
 GGCTTGGCAGCTGGGGAAACCCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC
 CCCACCCCTGTGTATTCTTGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG
 CGCTCCATACAGTTCTCAGAGCGGGTCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT
 CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC
 CCCACCCCTCAGACCCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT
 CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
 GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG
 GCGCTTGGCTGCTGGCCGCGCATCATCAGCTGGGCGGAGGCTGTGCCGAGCGCAACAGGCCC
 GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
 GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCGG
 CCGCGCGCTCT**AG**GGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGCGCGCCAGATCCACA
 TCTGGATCTGGATCTGCGGCGGGCTCGGGCGGTTTCCCCGCCGTAAATAGGCTCATCTACC
 TCTACCTCTGGGGGCCCGGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCCCGAC
 GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCAC
 GACTTCCGGCCCCGCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTAT
 AGGTATTGTGAACCTGCCACATATCTTATTATTCTCCTCCAATTTCAATAAATTATTTATT
 CTCCAAAAAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318
><subunit 1 of 1, 317 aa, 1 stop
><MW: 33732, pI: 7.90, NX(S/T): 1
MVVSGAPPALGGGCLGTFSTLLLLLASTAILNAARIPVPPACGKPKQLNRVVGGEDSTDSEWP
WIVSIQKNGTHHCAGSLTSTRWVITAAHCFKDNLNKPYLFVLLGAWQLGNPGSRSQKVGVA
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG
VPLPHPQTLQKLKVPFIIDSEVCSHLYWRGAGQGPIITEDMLCAGYLEGERDACLGDSGGPLMC
QVDGAWLLAGIISWGECAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
SGAAARS

Signal sequence:

amino acids 1-32

N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 99

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCTGTGCTTCTGCTGCCGCTACTGCTA
CTGCTGGTGGCCACCACAGGCCCGTTGGAGCCCTCACAGATGAGGAGAAAACGTTTGATGGT
GGAGCTGCACAACCTCTACCGGGCCAGGTATCCCCAGCGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCCTGTGGGGCCAC
AAACAAGGAGCGCGGGCGCCGCGCGAGAATCTGTTTCGCCATCACAGACGAGGGCATGGACGT
GCCGTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACCTCCCTCTGTGAACCCATCGGAAGC
CCGGAAGATGCTCAGGATTTCGCTTACCTGGTAACCTGAGGCCCCATCCTTCGGGGGACTGA
AGCATCAGACTCTAGGAAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAACCCAGGCC
CCAACTTCTCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCCTAAC
AACTGAGGTCCCCTTCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA
CCTTCCCCAAATCGACCCATGTTCTATCCAAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
GGAACCTCTACCCCATGCCAGGAGGAGGTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGGCGTCCCTGGCTCTGCAGTCGTCTTGCAGGTGCAGAGGGCCCTGACA
AGCCTAGCGTTGTGTACAGGGCTGAACCTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA
CTACTGCTCCTGCCTCCTCTGGTGTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG
TGAAGAGGTCAGCTGTCTCTGTCTATCTTCCCCACCTGTCCCCAGCCCTAAACAAGATA
CTTCTTGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC
ATCTGGAGGCACAAGCCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC
GGGCCACACCTCTCCTGCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT
CACTGCCTACCTGGCCTGGGGCTGTCTGCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG
TGTAGCTGGGGATGGGGATTCTTAGGGCAGATGAAGGACAAGCCCCACTGAGTGGGGTTC
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAACCT
GTCCAACCTGTGAAA

FIGURE 100

MHGSCSFLMLLLPLLLLLLVATTGPGVGTDEEKRLMVELHNLRYAQVSPITASDMLHMRWDEE
LAAFAKAYARQCVWGHNKERRRGENLFAITDEGMDVPLAMEEWHHEREHYNLAAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLGVEETNIELLCNYEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLPLYLVTEAPSPRATEASDSRKMGTTPSSLATGIPAFVLVTEV
SGSLATKALPAVETOAPTSLATKDPSPMATEAPPVCVTEVPSILAAHSLPSLDEEPTVPFKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEALPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGARGPDKPSV
SGLNSGPGHVWVPLGLLLLPLVLGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,
250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

FIGURE 101

GTAACTGAAGTCAGGCTTTTCATTGTGGGAAGCCCCCTCAACAGAATTGGTCATTCTCCAAGTTATGCGTGGACGT
 ACTTCTGTGTGTTCTCCCTCTGCTTGCTTTTTACATTAGCAGACCGGACTTAAGTCACACAGATTATCTTTCA
 CAGGCAAGTTCATGAGCCACCTTCAAAGCCTTGGAGAGTGAACCTGAACCAACAAATGAATTGGAGCCACTCC
 AAATCTGGGACCGACTCTCGGCAATATTACACTTCTCTCTTGGCTGGAAACAGGATGTTGAAATATCTCCCTGA
 ACATCTGAAAGAGTTTTCACTCCCTTGAACCTTTGGACCTTAGCAGCAACAAATATTTTCAGAGCTCCAAACTGCAAT
 TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGCGAACCTGGGTATTTTGACAA
 TTTTGGCCAAACACACTCTCTGTGTTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCACCCGAAGATGTTTAAACT
 GCCCAACTGCAACATCTCGAATTGAACCGAAACAGGATTAAAAATGTAGATGGACTGACATTTCCAAAGCCCTTGG
 TGCTCTGAAGTCTCTGAAATGCAAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTGGGGCTGGAGCA
 CATTGAAATTTTTCAGCTGGACCTAAACAACTTAACAGAGATTACCAAAAGGCTGGCTTTACGGCTGCTGATGCT
 GCAGGAATCTCATCTCAGCCAAATGCCATCAACAGGATCAGCCCTGATGCTTCCGGGGCTTCCAGTTTAAAGACTT
 TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATAC
 ACTGCACATTGGGAACCAACAGCTCAGCTACATTGCTGATTGTCCTTCCGGGGCTTCCAGTTTAAAGACTT
 GGATCTGAAGAACAAATGAAATTTCTGGACTATTGAAGCATGAATGGTGTCTTCTCTGGCTTGACAACTGAG
 CGGACTGATACTCCAAAGAAATCGGATCGCTTCTATTATAAAAGGCTTCACTGGTTTGGATGCATTGGAGCA
 TCTAGACCTGAGTGACAACGCAATCATGCTCTTACAGGCAATGCAATTTTCAAAATGAAGAACTGCAACAAAT
 GCATTTAAATCAACTAAGCCCTTTTGGCATTCAGACTAAATGGCTCCACAGCTGGGTGGCGGAAACAACTT
 TCAAGACTTTGTAATGGCCAGTTGTGGCCATCTCAGCTCTCAAAAGGAGAGCACTTTTGCCTGTAGCCGAGA
 TGCTTTGTGTGATGATTTTCCCAAAACCCAGATCAGGTTTCAGCCGAGAACACAGTCCGCAATAAAGGTTT
 CAATTTGAGTTTCATCTGCTCAGCTGCGAGCAGTGAATTCCTCAATGACTTTTGGTTGAAAAAGCAATGA
 ACTACTGCATGATGCTGAAATGGAAATTTATGCACACTCCGGGCCCAAGGTGGCAGAGTGATGAGTATATCCAC
 CATCCTTCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGCTATCTCAATCACTTTGGTTC
 ATCCTACTCTGTCAAAGCCAGCTTACAGTAAATATGCTTCCCTCAATTCACCAAGACCCCATGGATCTCAACAT
 CCGAGCTGGGGCATGCGACGCTTGGAGTGTGCTGCTTGGGGCACCCAGCCCCGAGATAGCTGGCAGAGGA
 TGGGGGCACAGACTTCCAGCTGCACGGGAGAGACGATGCATGTGATGCGCCGAGGATGACGTGTTCTTTATCGT
 GGATGTGAAGATAGAGGACATTTGGGGTATACAGCTGCACAGCTCAGAACTGCAGGAAGTATTTACGAAATGC
 AACTCTGACTGCTCTAGAAACACATCATTTTTTGGGCCACTGTGGACCGAAGCTTAACCAAGGGAGAAACAGC
 CGTCTCAAGTGCATTTGCTGGAGGAAGCCCTCCCCATAACTGAAGTGAACCAAGATGATAGGCCATTTGGTGGT
 AACCGAGAGGCACTTTTTCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAAGTGTGCTGGGAA
 ATACACATGTGAGATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTCGCTCAGTGTGATCCCACTCCAAC
 CTGCGACTCCCTCAGATGACAGCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGCTGATCATAGC
 CGTGGTTGTGCTGTGTGGGGCAGTCACTGTGTGGGTGGTCAATATACCAACAGGCGGAGGAATGAAGA
 TTGCAGCATTACCAACACAGATGAGACCAACTTGCAGCAGATATTCCTAGTTATTTGTCACTCTCAGGGAACGTT
 AGCTGACAGGCAAGGATGGGTCTGTCTTCAGAAAGTGGAGCCACCAAGCTTTGTCACTCTCAGGTGCTGG
 ATTTTCTTACCACCAATGACAGTGTGGGACCTGCCATATTGACATAGCAGTGAAGCTGATGTGGAAGCTGC
 CACAGATCTGTGTTCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGAAATGTGTATGGCTCAGA
 TCCTTTTGAACATATCATACAGGTTGCACTCTGACCCAGAAACAGTTTAAATGGACCACTATGAGCCGAGTTA
 CATAAAGAAAAGGAGTGTCATCCATGTTCTCATCCTTCAGAAAGATCCTGCGAAGCAGGCTTCAGTAATATATC
 GTGGCCCTCAGATCTGAGGAAGCTACITTAACACTAGTTACTCTCAATGAAGGAGCTGGAATGAAAAATCTGTG
 TCTAAACAGGCTCCCTCTTGTAGATTTTGTGCAAAATCCAGAGCCAGCGTGGTGGTCCCTCAGTAAATCTTTTCA
 TACCTTTGGAAGGCTCTCAGGAGACTCACCTAGATGCTTATCAAGCTTTGGACAGGCTCAGATTGTGACGCC
 AAGAGCCCTTTTATTTGAAGCTCATTTCTTCCAGACTTTGGACTCTGGGTGACGAGGAATGGGAAAGGAGAC
 AGATTTTCAGGAAGAAATACATTTGTACTTTAAACAGACTTTAGAAACTACAGGACTCCAAATTTTCAGTC
 TTATGACTTGGACACATAGACTGAATGAGACCAAGGAAAGCTTAACATACTCTCAAGTGAACCTTTTATTA
 AAAGAGAGAGAATCTTATGTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAATGCTTTATTTATACAGAT
 GAACCAAAATGCAAAAAGCTATGAAATTTTATACCTGGGATGATGCTCATATAAGAAATACCTTTTAAACTA
 TTTTTTAACCTTTTATTTGCAAAAAGTATCTTACGTAATTAATGATATAAATCATGATTTATTTATGATTT
 TTTAATGCGAGATTTCTTTTATGAAAAATGAGTTACTAAAGCATTTTAAATTAATACCTGCTGTGACCTTTT
 TAAATAGAAATTTCTTATTTATTTTGCACATTTATTTAATAAATGTGATTTGAA

FIGURE 102

MVDVLLLFSLCLLPHISRDPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVVSAN
ITLLSLAGNRIVEILPEHLKEFQSLLETLDLSSNNISELQTAFPALQLKYLVLNSNRVTSMEP
GYFDNLANTLLVLKLNRRNRIASAIPPKMFKLPLQHLLELNRRNKIKNVDTGTFQGLGALKSLKM
QRNGVTKLMDGAFWGLSNMEILQLDHNNTETITKGWLYGLMLQELHLSQNAINRIISPDWAE
FCQKLSLELDLTNNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSSLKTLDLKNEE
ISWTIEDMNGAFSGGLDKLRLILQGNRIIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCDCQLKWLQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCD
DFPKPQITVQPTQSAIKGSNLSFICSAASSSDSPMTFAWKDNELLHDAEMENYAHLRAG
GEVMEYTTILRLREVEFASSEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA
MARLECAAVGHPAPQIAWQKDGDTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTVLETPSFLRPLLDRTVTVKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER
HFFAAGNQLLIIVSDSDVSDAGKYTCESMNTLGTGERGNVRLSVIPTCTCDSPQMTAPSLDDDG
WATVGVVIIAVVCCVVGTSVLVWVVIYHTRRRNEDCSITNTDETNPADIPSYLSSQGTAD
RQDGYVSSSESGHHQFVTSSGAGFFLPQHDSSGTCCHIDNSSADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDFPETYHTGCSPPDPTVLMHDHYEPSYIKKKCYPCSHPSSESCERSFSNLSW
PSHVRKLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMCTFGKALRPHLDA
YSSPGQPDGCPQAFYHLKAHSSPDLDSSGEEDGKERTDFQENHICTFKQKTLENYRTPNFQS
YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 103

GGGGAGAGGAATTGACCATGTAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAATG
AAGGATG CAGGACG CAGCTTTCTCTCGGACCGAACGCAATGGATAAACTGATTGTG CAGAGAGAGAGGAAGAAC
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA
TAAAC CAGAGT TAGACCCGCGGGGTTGGTGTGTTCTGCACATAAATAAATAATCTTTAAAGCAGCTGTTCCCTCCC
CCACCCCAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAGTATGTTTCATTTTTTCTC
TATAAAGAGAGAAATGAGCCAGGAGATATTTTTTGAATGAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT
GGTGTGGTGGTGTTTTTCTTTCTTTTGAATTTCCACAGAGGAGAGAGAAATTAATAATACATCTGCAAGAAAA
TTTCAGAGAGAGAAAGTTGACCGCGGAGATTGAGGCATTGATTGGGGAGAGAAACAGCAGCAGCAGTGTGA
TTTTGGCTATGTTGACTAAAAATTGACGGATAATTGCAAGTGGATTTTTCTTCATCAACCTCTTTTTTTAAAT
GTTTATCTCTTTGGTATCAAGATCATGCGTTTTCTCTGTCTTAAACACCTGGATTTCATCTGGATGTTGCT
GTGATCAGTCTGAATAACAAGTGTGAAATCCAGAGGACCAACACAGATAAATTAATGAATGTTGAACAAGAT
GACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAAACAGGGCCCTATTGACCCCTGCTTGTGTGCT
GCTGGCTCTCAACTTCTTGTGGTGGCTGGTCTGGTGGCGGCTCAGACCTGCCCTCTGTGTGCTCTCTGCAGCAA
CCAGTT CAGCAGGTGATTGTGTTCCGAAAAAACCCTGCGTGGGTTCCGGATGGCATCTCCACCAACACACGGCT
GCTGAACCTCCATGAGAACCAAAATCCAGATCATCAAAGTGAACAGCTTCAAGCATTGAGGCATCTGGAAATCCT
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATGGGGCTTCAATGGTCTGGCGAACCTCAACACTCTGGA
ACTCTTTGACATCTCTTACTACCATCCGAACTGGAGCTTTTGTATCTGTCTAAACTGAAGGAGCTCTGGTT
GCGAAACACCCCATTTGAAAGCATCCCTCTTTATGCTTTTAAACAGAATTCCTCTCTTGCGCCGCTAGACTTAGG
GGAAATTGAAAGACTTTTCATACATCTCAGAAAGTGCCCTTTGAAGGTCTGTGTCACACTTGAGGATTTTGAACCTTG
CATGTGCAACCTTCGGGAAACCTTACACTCAGCCGCTCATAAACTAGATGAGCTGGATCTTTCTGGGAATCA
TTTATCTGCCATCAGGCTTGGCTCTTTCCAGGGTTTGATGACCTTCAAAAATCTGGATGATACAGTCCCCAGAT
TCAAGTGATTGAACGGAAATGCTTTGACAACTTCAGTCACTAGTGGAGATCAACTGGCAGCAACAATAATCTAAC
ATTA CTGCTCATGACCTCTTCACTCCCTTGATCATCTAGAGCGGATACATTTCATCAACAACCTTTGGAAGT
TAACGTGTGACATACCTGTGGCTCAGCTGGTGGATAAAGACATGGCCCTCGAACACAGCTTTGTGCCCCGGTG
TAACACTCTCTCCAACTCTAAGGGGAGGTACATTGGAGAGCTCGACAGAAATTAATTCATGCTATGCTCCGGT
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGATGGCAGCTGAGCTGAAATGTGCGGCCCTCCACATC
CCTGACATCTGATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTCCGGATAGCTGT
GCTCAGTGATGGTACGTAAATTTCACAAATGTAACGTGCAAGATA CAGGCATGTACACATGTATGGTGGATAA
TTCCGTTGGGAATACTACTGCTTCAGCCACCTGAAATGTTACTGCAGCAACCACTACTCCTTTCTTACTTTTC
AACCCTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGAGACACAGATAACAATGTGGGTCCCCTCC
AGTGGTCTGACTGGGAGACCAACCAATGTGACCACTCTCTCACACCAAGAGCAAGGTCGACAGAGAAAAACCTT
CACCATCCCACTGATCTGATATAAACAGTGGGATCCAGGAATTGATGAGGTCTAGAGACTACCAAAATCATCAT
TGGGTGTTTGTGGCCATCCACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA
TCGGCAAAACCATCAGCCGCCAACAGGACTGTGAAATTAATAATGTGGATGAGATTACGGGAGACACACC
CATGGAAGGCCACTGCCATGCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT
CAACCACACAACACAGTTAACACAAATAAATTAATACACAGTTCAAGTCATGACCGTTATTGATCCGAATGAA
CTCTAAAGAGATTACAAGAGACTCAAACTTAAACATTACAGAGTTACAAAAAACCAACAAATCAAAAAAAA
GACAGTTTATTAATAATGACAAATGACTGGCTAAATCTACTGTTTCAAAAAAGTGCTTTACAAAAAAAACAA
AAAAAGAAAGAAATTTATTTATTAATAATCTTATTTGTGATCTAAGCAGACAAAA

FIGURE 104

MLNKM TLHPQQIMIGPRFN RALFD PLLVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGI STNTRLNLN LHENQIQI IKVNSFKHLRHL EILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLRBIPNLTPLIKDELDSLGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVR IAVLSDGT LNFNTNVTVDQTMGYTCMVSN SVGN
TTASATLNVTAATTPFSYFSTVTVETMEPSQDEARTDNNVGPTPVVDWETTNTVTTSLTPQ
STRSTEKTF TIPVTDINSGIPGIDEVMKTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHNYSYKSPFNHTTNTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,
442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

FIGURE 105

AGCCGAGCTCTCAAGCTGCAACTCTGTGTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCTGCTGTTTGGGGGCA
TGAAAGGGCTTCGCCGCCGGAGTAAAGAAGAAATTGACCGGCGAGCGGAGGAGGAGCGCGACGCGGACCGC
GAGGGCGCGCTGCACTCCCTCGGCTGGAAAGTTGTGCGCGGGCCCGAGCGCGCGCGCTGGAGAGTAC
GACTAGGCGCGCTGGACCGCTGCGCGCGCGCGGAGCTCCGTGCGCGCGCGCGGGGTGGGGGTGCTGCTGTGTC
GCGGTCTGGGGCGCGCTGCGCTGCGTCCGACAGCGCGCGCTGCGGGGAACCTTTTGACAACTTTGGCAGCAAC
GAGCGCCCATGCCCACTACCTTGCCGCTGCTCGGGGACCTTGCTGGACTGACAGTCGTAAGCGGGTAGCGCGCTCT
CCGCGCAACTCCCGCTCTGGCTCGGCTGGAGCTTAAGTACAAACAGATTATCTTTCTATCAAGGCAAGTTCC
ATGAGCCACCTTCAAAGCTTCGAGAAGTGAACATGAACAACAAATGAATTGGAGACCAATCCAAATCTGGGACCA
GTCTCGGCAAAATATTAACATTCTCTCTCTGGCTGGAAACAGGATTGTTGAAATCTCCCTGAACATCTGAAAGAG
TTTCAGTCCCTTGAACCTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAACTGCATTTCAGCGCTACAG
CTCAAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACTCCGTTATTTTGACAACTTTGGCACAACACA
CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCACCCCAAGATGTTTAAACTGCCCACTGCAA
CATCTCGAATTGAACCGAAACAGAGATTAAAAATGTAGATGGACTGACATTCCAAGGCTTGGTGCTCTGAAGTCT
CTGAAAATGCAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTGGGGGCTGAGCAACATGGAATTTTG
CAGCTGGACCATTAACAACCTTAACAGAGATTACCAAGGCTGGCTTTACGGCTTGCTGATGCTGACGAGAACTTCAT
CTCAGCCAAAATGGCCATCAACAGGATCAGCCCTGATGCTCGGGAGTTCTGCCAGAAGCTCAGTGAGCTTGGACCTA
ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGCGCTAAAGCTTACTAAATACATGACATCTGGG
AACACAGAGTGCAGCTACATTGCTGATTGTCCTTCCGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAGAAAC
AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTTCTCGGGCTTGACAAACTGAGGGGACGTAGACTC
CAAGGAAATCGGATCGGCTCTATTACTTAAAGAAAGCTTCACTGGTTTGGATGATCTGGAGCATCTAGACCTGAT
GACACGCAATCATGCTTTTCAAGGCAATGCTTTTCAACAATGAAGAACTGCAACATTCGATTTTAAATACA
TCAAGCTTTTGTGCGATTTGCCAGCTAAAATGGCTCCCAACAGTGGGTGGCGGAAACAACTTTCAGAGCTTTGTA
AATGCCAGTTGTGCCATCTCAGCTGCTTAAAGGAAGAAGCACTTTTGGCTTAGCCAGATGGCTTTGTGTGT
GATGATTTTCCCAAACCCAGATCAGCGTTACGACCAAGAACACAGTCCGCAATAAAGGTTCCAAATTTGAGTTTC
ATCTGCTCAGCTGCCAGCAGCAGTGAATTCCTCAATGACTTTTGTCTGGAAAAAGACATGAACTACTGCACTGAT
GCTGAAATGGAAAATATGACACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAACCTCTCGGCTG
CGCGAGGTGGAATTTGCGCAGTGAGGGGAAATATCAGTGTGTCATCTCCAACTCACTTTGGTTCTATCTCTGTC
AAAGCCAAAGCTTACGAAATATGCTTCCCTCATTCAACCAAGACCCCAATGACTTCCACCTCCAGCTGGGGCC
ATGGCAGCTTGGAGTGTGCTGCTGTGGGGCACCAGCCCCCAGATAGCTTGGCAGAAGGATGGGGCGACAGAC
TTCACAGCTGCACGGGAGAGCGCATGCAATGTGATGCCGAGGATGACGTGCTTTTATCGTGGAATGTGAAGATA
GAGGACATTTGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTACGACAAATGCAACTCTGACTGTC
CTAGAAAACACTCAATTTTGCAGGCACTGTTGGACCGAAGCTGTAACCAAGGAGAAACAGCCGCTCTACAGTGC
ATTGCTGGAGGAAGCCCTCCCCCTAACTGAATGGACCAAGAGATGATAGCCCAATTTGGTGGTAACCGAGAGGAC
TTTTTTGAGCAGGCAATCAGCTTCTGATATTGTGGACTCAGATGTGATGATGTCGGGAAATAACATGTTGAG
ATGCTTAACACCTTTGGCACTGAGAGAGGAAACGTGGCGCTCAGTGTGATCCCACTCAACCTGCGACTCCCT
CAGATGACAGCCCATCTGTTAGACGATGACGGATGGGCACTGTGGGTGCTGATCATAGCCGCTGTTGTGT
GTGTTGGGCGCTCCTGTTGGGTGGTCATATATACCACAAAGCGGAGGAAATGAAGATTGCGAGCATACC
AACACAGATGAGACCAACTTGCACAGCATATTCCTAGTTATTGTCTATCTGAGGAAAGCTTGCACAGGACG
GATGGGTACGTGCTTCAGAAAGTGAAGCCACCACTGTTGTCACTTTCAGGTGCTGGATTTTCTTACCA
CAACATGACAGTATGGGACCTGCCATATTGACAAATGACAGTGAAGCTGATGTGGAGAACTGCCACAGATCTGTC
CTTTGCTCGCTTTTGGGATCCACAGGCCCTATGATTTTGAAGGAAATGTGATGGCTCAGATCCTTTTGAACA
TATCATCAGTTGCGAGCTCTGACCCAAAGAACAGTTTAAATGGACACTATGAGGCCAGTATACATAAGAAAAAG
GAGTGTACCTACCTATCTCATCTTTCAGAAAGATCCTGCGAACGAGAGCTCAGTAATATATCTGCGGCTTCACAT
GTGAGGAAGCTACTTAACTAGTTACTTCTCAAAATGAAGGACCTGGAATGAAAATCTGTGTCTAAACAGTGC
TCTTTAGATTTTATGCAAAATCCAGAGCCAGCGCTCGGTTGCCCTCAGTAATTTCTTTCATGGGTACCTTTGGAAAA
GCTCTCAGAGACCTCACCATGATGCTTATCAAGCTTTGGACAGCCATGAGCTCAGCCCAAGAGCTTTTAT
TTGAAAGCTCACTTTTCCCAAGCTTGGACTTGGGTGAGAGGAAGTGGAAAAAGAGGACAGATTTTTCAGAA
ACATAGACTGTAATGAGACCAAGGAAGCTTAAACATACTACCTCAAGTGAACCTTTTATTTAAAGAGAGAAAT
CTTATGTTTAAATGGAGTTATGAATTTTAAAGGATAAAATGCTTTTATACAGATGAACCAAAATATAC
AAAAGTATATGAAATTTTATCTGGGAATGATGCTCATATAAGAAATACCTTTTAAACTATTTTTTAACTTTG
TTTTATGCAAAAAGTATCTACGTAATAATGATATAAATCATGATATTTTATGTTATTTTATATGCCAGA
TTTCTTTTATGGAATAGTACTAAAGCAATTTTAAATAATACCTGCTTTGACCACTTTTATTAATGGAAGTT
ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGCAATTTGAAAAAAGAAAAAAGAAAAAAGAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 106

MSAPSRLRAAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCGLDGLLDCSR
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT
LLSLAGNRIVEILPEHLKEFQSLLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRRRI SAIPPKMFKLPLQLHLELNRNKIKNVDLGTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC
QKLSLDDLTFNHLRSRLDDSSFLGLSLNLTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS
WTIEDMNGAFSGLDKLRRLILQGNRIRISITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNLTSSLLCDCQLKWLPLQWVAENNFQSFVNASCAPQLLKGRSIFAVSPDGFCVDDF
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKDNELLHDAEMENYAHLRAGGGE
VMEYTTILRLREVEFASGKYQCVISNHFSSSYSVKAKLTVNMLPSFTKTPTMDLTI RAGAMA
RLECAAVGHAPQAIQAWQKDGDTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCAQNSA
GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF
FAAGNQLLIIVDSDVSDAGKYTCESNLTGTERGNVRLSVIPTPTCDS PQMTAPSLDDDDGWA
TVGVVILAVVCCVGTSLVWVVIYHTRRRNEDCSITNTDETNLPAIDPSYLSSQGT LADRQ
DGYVSSSESGSHHQFVTSSGAGFFLPQHDSGCTCHIDNSSEADVEATDLFLCPFLGSTGPMY
LKGNVYGSDDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCHSPSEESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDAYS
SFGQPSDCQPRAFYLKAHSSPDLDGSGSEEDGKERTDFQEENHICTFKQTLNRYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

FIGURE 107

CAAAACTTTCGCTCGCGGAGAGCGCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC
TGGGGGAGCGCGTTTCGGCTTGTGGGGCGCCCTCGGCGCCGGGGCGCAGCAGGAGGAGGGGAAGCTGTGGTCTGCC
CTGCTCCACGAGGCGCCACTGTGTGAAACCGGAGAGCCCTTGGTGGTCCCGTCCCTATCCTCTTTATATA
GAAACCTTCCACACTGGGAAGGAGCGCGCGAGGCGAGGAGGGCTCATGGTGAGCAAGGAGGCGCGGTGATCTGCAG
GCGCACAGCATTCGAGATTACAGATTTTACAGATACCAATGGAAGGCGAGGAGGCGACAGAGCTTCCTGGT
TCCATCAGCCCTGGCGCCAGGCGCATCTGACTCGGCACCCCTGCAGGACCATATGCCCCAGAGCCGGGTGCTGC
TGCTCTGTGCTGTGCGGCCACAGCTGCACCTGGGACCTGTGCTTGCCTGTGAGGCGCCAGGATTTGGCCGAA
GTGGCGGCCACGCTGAGCCCCGAAGAGAAAGAAATTTGCGGAGGAGGAGCGCGGTGCTGTACTGAGCCCTGAGG
AGCCCCGGGCTGGCCGAGCGCGGTGAGCTGCCCCGAGACTGTGCTTCCAGGAGGCGGTCTGTGACTGTG
GGGTAAATTGACTGTGCTGAGTTCCCGGGGAGCTGCTGAGCACACCAACCACTATCTCTGCAGAACCAACCAGC
TGGAAAGATCTACCTCGAGAGCTCTCCCGGCTGCACCGGTGAGACACTGTCAAGCTGCAAAACACACCGCTGA
CTTCCCGAGGGCTCCAGAGAAAGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC
TGACCTTGGCACCCCGCTTCTGCGCAACGCCCTGATCAGTGTGGACTTTGCTGCCAATATCTCACCAGATCT
ATGGCTCACTTGTGCCAGAAAGCAAACTTGAGGTCTGTGTACTGTGCACAAACAAGCTGGCAGACCGCGGGC
TGCCGAGCAACATGTTCAACGGCTCCAGCAACGTCGAGGTCTCATCTGTCTGCAGCAACTTCTGCGCCACAGTGC
CCAAGCACTGCGCGCTCGCTGTACAAGCTGCACCTCAAGAACACAAAGCTGGAGAAGATCCCCCGGGGGCTC
TCAGCGAGCTGAGCAGCTTGGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGCGCTGGACAAACGAGA
CTTTCTGGAAGCTCTCCAGCGTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCCACTGGGCTGC
CGCGCAGCTGGTGTGCTGTCACTTGGAGAAAGAACGCCATCCGAGCGTGGACGGGCAATGTGCTGACCCCCATCC
CGAGCTGGAGTACTGTGCTGTGCACAGCAACAGCTGCGGGAGCAGGGCATCCACCACTGSGCTTCCAGGGCC
TCAAGCGGTGACACGGTGCACCTGTACAACCAACCGGCTGAGGCGGGTGCACCACTGSGCTTCCAGGGCTGC
GACCCCTCATGATCTTGCACACACAGATCAAGGCATTGGCGCGAAGACTTGGCCACCACTACTTCTCGGAGG
AGCTCAACCTCAGTACAACCGCATCACCAGCCACAGGTCACCGGAGCGCTTCCCAAGACTGCGCTGCTGAGG
GCTCTGGAGCTGTGGGCAACCGGCTGCACAGCTGCCACTGGCTGCCTCGAAATGTCACTGTGCTGAAGG
TCAAGCGCAATGAGCTGGCTGCCTTGGCACAGGGGGCGCTGCGGGCATGSGCTCAGCTGCGTGAGCTGTACTCA
CCAGCAACCGACTGCGCAGCGAGCCTGGGCCCCGTGCTGGGTGGACTCGCCCATCTGCAGCTGCTGGACA
TGCCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCGAGTCACTTGAGTACCTGTACTCGAGAACACA
AGATTAGTGGGTGCCCGCAATGCCTTGCCTCCAGGCCCACTCAAGGGGATCTTCTCAGGTTTAAACAGC
TGGCTGTGGGCTCGGTGTGGACAGTGCCTCCGAGGCTGAAGCACTCGAGGTCTTGGACATTGAAGGCAACT
TAGAGTTTGGTGACATTTCCAAAGACCGTGCGCGCTTGGGGAAGGAAAGGAGGAGGAGGAGGAGGAGGAGG
AGGAAGAGGAAACAAGATAGTGACAAGGTGATGCAGATGTGACCTAGGATGATGGAACCGCGGACTCTTTCTGC
AGCACACGCTGTGTGCTGTGAGCCCCCACTCTGCGGTGTGCACAGACACACCCAGCTGCACATGAGGCA
TCCCACATGACACGGCTGACACAGTCTCATATCCCCACCCCTTCCACGCGGTGTCCACGCGGCAGACATGC
ACACACATGACACCCCTCAAACACCCAGCTCAGCCACACACAACCTACCTCAAACACCAACAGTCTCTGTACAC
CCCCACTACGCTGCCACGCCCTTGAATCATGCAGGGAAGGGTCTGCCCTGCCCTGGCACACAGGCAACCA
TTCCTTCCCCCTGCTGACATGTGATGCGTATGCATACACACACACATGCACAGTCAATGTGTGGAA
CAGCCCTCCAAAGCTGATGCCACAGACGCTCTTGGCCCCAGCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT
GTCCATCTGTCCGTCGTTCCCTGGAGAAGACACAAGGATCCATGCTCTGTGGCCAGGTGCTGCGCACCTCT
GGAATGCACAAAGAGCTGCTTTATCTCTTCCCCATCTATGGGGACAGGAGGCTCAGGATCTGCTGGCTGGCC
TGGCCCAACCTGCTCTCCGCTGCTGGGAGTCACTCTGCTAAGAGTCCCTCCCTGCCAGCCCTGGCAGGACA
CAGGCACTTTCCAAATGGGCAAGGCCATGTGAGGCGAGGTGGGAGAGCCCCCTGGGTGCTGCTGGGGCTTGGG
CAGGAGTGAAGCAGAGGTGATGGGCTGTGGCTGAGCGCAGGGAGGAGGACCTGCATCAGGAGCACTTT
GTCTTTCAGGCGCTTGGGGAGTTCGGGTGCCTTTATTTTTTTATTCTTTTCAAGGAAAAAATGATAAAAAT
CTCAAGCTGATTTTTCTGTATAGAAAACTAATAAAGCATTATCCCTATCCCTGCAAAAAATAA

FIGURE 108

MEGEEAEQPAWFHQPWPRGASDSAPPAGTMAQSRVLLLLLLLLLPQLHLGPVLAVRAPGFGRS
GGHSLSPREENEFAEEEFVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCCGIDLREFFGDLF
EHTNHLSLQNNQLEKIYPEELSRHLRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLHSN
QLREQGIHPLAQGLKRLHTVHLYNNALERVPSGLPRRVRTLMIHNNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKIGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEEEETR

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

FIGURE 109

GGGAGGGGGCTCCGGGGCCGCGCAGCAGACTGTCTCGGCGCCGCGCCTCGCCGCTGTCTCTCGGGAGCGGCAG
CAGTAGCCCGGGCGGGAGGGCTGGGGGTTTCCTCGAGACTTCTCAGAGGGGGCGCCTCCATCGGGCCACACCC
CAACCTGTTCCTCGCGCCCACTCGCTGCGCCCAAGAACCCGCTGCGCAAC**TG**GATTTCCTCGGGCTGGT
CTGTGGTATCTCTCGCTACTCTGCAAGCGCGCCCGCGAGTTGCAGCGGAGGTGGCCCAAGCAATAGTGTATCGAT
TGGCCTATGTGCTTATGGTGGGAGGATTGACTGCTGCTGCGGCTGGGCTCGCCAGTCTTGGGAGCAGGTGTCAGCC
TGTGTGCCCAACACGATGCAACATGTGTAAATGTATCGGGCCAAACAAGTGCAGGTGTCACTCTGTTATGTCTGG
AAAAACCTGTGAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCGGCCCTGTAAAGCAGAGTGCATGAAGCTTAC
CGCAGCTACAAAGTGTACTGTCTCAACGGATATATGCTCATGCCGATGGTTCTGTCTCAAGTGGCCCTGACCTG
CTCCATGGCAAACTGTGAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCTGGCCT
GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCCTCTGCTCATAGATT
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCTAAAGGCTTCGATCTCATGTATGTATTGGAGG
CAAATATCAATGTCAATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA
CGTACGTGGCTCTCAAGTGCAAATGTAAAGAAGGATACAGGGTGTGGAAGTGTGCTGTGTATATCCCAAA
AGTTATGATTGAACCTTCAGGTCCAATTCAATGTACCAAAAGGAAATGGTACCAATTTAAAGGGTGACACAGGAAA
TAATAATTTGGATTCCTGATGTTTGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCTATCATTTACCA
CAGCCTACTTCTAAGCCAAACAAAGACTACACCAAGCCAAACCAATTCCTACTCCACCACACACACAC
CCTGCCAACAGACTCTCAGAACACCTCTACCCTTACAACCCAGAAAGCCAAACACCGACTGACAACTATAGC
ACCAGCTGCCAGTACCTCCAGGAGGGATTACAGTTGACAAACAGGGTACAGACAGACCTTCAGAAACCCAGG
AGATGTGTTTCAGTGTCTGGTACACAGTTGTAAATTTGACCATGGACTTTGTGATGGATCAGGGAGAAAGCAA
TGACTTGCACCTGGGAACCAATCAGGGACCCAGAGGTGGACATATCTGACAGTGTGCGCAGCCAAAGCCCAGG
GGGAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCTCATGCACTCAGGGACCTGTGGCTGTCTATTCAAGCA
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAATCAAGTGGCCACAGGAGCAGCCCTGTG
GGGAGAAATGGTGCCATGGCTGGAGGCAACACAGATCACCTTGCAGGGGCTGACATCAAGAGCGAATCACA
AAGAT**GA**TTAAAGGGTTGAAAAAAAGATCTATGATGGAATAATTAAGGAATGGGATTATTGAGCCTGGAGAG
AGAAGACTGAGGGGCAAAACATTTGATGGTTTCAAGTATATGAAGGGTTGGCACAGAGAGGTGGCGACAGCTG
TTCTCCATATGCACATAAGAATAGAACAGAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGCAGG
GGCATTGTGATAATCTTCTAAAAAAAGAGGTGAAATCTCAGTATCTCTCTCTTCTTAAAAAAATTAGA
TAAAAATTTGCTATTTAAGATGGTTAAAGATGTTCTTACCAGGAAAGTAAACAAATTAAGAATTTCCCAAA
AGATGTTTGTATCTACTAGTAGTATGCAGTGAAATCTTTAGAACTAAATAATTTGGACAGGCTTAAATTTAGG
CATTTCCCTCTTGACCTCTTAATGGAGAGGGATTGAAAGGGGAAGAGCCCAACAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAGAAAAAAGGAAACTATTTATTCCAAATGAGAGTATGATGGAC
AGATATTTTATGATCTCAGTAAATGCTAGTGTGGCGGTGGTTTTCAATGTTTCTCATGGTAAAGGTATAAGCC
TTTCATTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTAAAGAGATCCTTCAAGGAACACAGTTTCAAGAG
ATTTTTCATCGGGTGCAATCTCTCTGCTCTGCTGTGTGACAAGTTACTTGGCTGCTGAGAAAGAGTGCCCTGGCCC
ACACCGGCAGACCTTTCTTCACTCATCAGTATGATTCAAGTTTCTTATCAATTTGAGCTCTCCAGGTTCCAC
AGAACAGTAATATTTTTTGAACAATAGGTACATAGAAAGGTCTTCTGTCAATTTAATCTTGGTAAAGAGGCTGG
AGGGGGAAAAATAATCATTAGCCTTTTGAATACGGCAGAATATATGGCTGTAGATCCATTTTAAATGGTTTCAAT
TCCTTTTGTGTCATATACTGACAGCTGAAGTGAAGGGGAAATAATGAAATTTTACTTTTCTGATGCCAA
TGATACATTGCACTAAACTGATGGAAGAGTTATCCAAAGTACTGTATAACATCTTGTTTATTTTAAATGTTT
CTAAATAAAAATGTTAGTGGTTTTCCAAATGGCCTAATAAAAACAATATTATTGTAATAAAAAACATGTTAGTAAAT

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 110

MDFLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP
RCKHGE CIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCOYGCDVVKGQIRQCQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQC
VNTFGSYICKCHKGFGLMYIGGKYQCHDIDEC SLGQYQCSSFARCYNVRGSYKCKCKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKNGTILKGD TGNNNWIPDVGSTWWPPKTPYIPPIITNRP
TSKPTTRTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTGLTTIAPAASTPPGGITVDN
RVQTD PQKPRGDFVSVLVHSCNFDHGLCGWIREKDN DLHWEPIRDPAGGOYLTVSAAKAPGG
KAARLVLPGLRLMHSGDLCLCSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHWGRQTQI
TLRGADIKSESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTGGCCCCCTTAGATTGTGA
AATGTGGCTCAAGGCTTCACAACCTTTCTTTCTTTGCAACAGGTGCTTGCCTGGGGCTGA
 AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGCTC
 CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
 CACAATGCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAAATACC
 AACCAAGTTTACCATGATGCCACCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCT
 GATGAAGGCAATTACATCGTGAAGTCAACATTCAGGGAATGGAACCTATCTGCCAGTCA
 GAAGATACAAGTCACGGTTGATGATCCTGTGCACAAAGCCAGTGGTGCAGATTATCCTCCCT
 CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAAGGGGCACTCGG
 CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
 TCCCCAAAACAATACCTTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
 GCCTGGTGAGGAACCTGTGTCAGTGAAATGGAAGTGATATCATTATGCCCATCATATATTAT
 GGACCTTATGGACTTCAAGTGAAATCTGATAAAGGCTAAAAAGTAGGGGAAGTGTTTACTGT
 TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT
 CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAAGCATGGGCCTCGCTTAGAAGTT
 GCATCTGAGAAAGTAGCCAGAAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC
 CGGCAGGCAAGATGAAACTCATTTACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
 CACAGAAAGGAAAATCATTGTACCTTTAGCAAGTATAACTGGAATATCACTATTTTGTATT
 ATATCCATGTGTCTTCTCTTCTATGSAAAAAATATCAACCCTACAAGTTATAAAACAGAA
 ATAGAAGGCAGGCAGAAACAGAAATACAGGAAAGCTCAAAACATTTTCAGGCCATGAAGATG
 CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTCACAGG
 ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTGACACGTACAGT
 GTATGAAGTTATTGACACATCCCTGCCAGCAGCAAGACCATCCAGAGT**TGA**ACTTTTCATGG
 GCTAAACAGTACATTGAGTGAAATCTGAAGAAAACATTTTAAGGAAAAACAGTGAAAAAGT
 ATATTAATCTGGAATCAGTGAAGAAACAGGACCAACACCTCTTACTCATTTTCTTTTACA
 TGCAGAATAGAGGCATTATGCAAAATTGAAGTGCAGGTTTTTCAGCATATACACAATGTCTT
 GTGCAACAGAAAAACATGTTGGGGAATATTCCTCAGTGAGAGTCTGTTCTCATGCTGACGG
 GGAGAACGAAAGTGACAGGGGTTTCTCATAAGTTTTGTATGAATATCTCTACAAACCTCA
 ATTAGTTTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
 TGTGGAACCTTTACATTGTTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTATTAGTG
 TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTATTGTACAA
 CAAAGTAATAAGGATGGTTGTCAAAAAACAAACTATGCCTTCTCTTTTTTTTCAATCACC
 AGTAGTATTTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAGTCTTATTTTTA
 TTTTTTTTCAAGGAAAGATGGATTCAATAAATATTCTGTTTTGCTTTTAAAAA

FIGURE 112

MWLKVF¹TF²FLSFATGACSGLK³VTVP⁴SH⁵TVHGV⁶RQALYLPV⁷HYGF⁸HTPAS⁹DIQ¹⁰I¹¹WL¹²FER¹³PH¹⁴
TMPKYL¹⁵LGSV¹⁶NKSV¹⁷VPD¹⁸LEYQ¹⁹HKFT²⁰MMPP²¹NASLL²²INPL²³QFP²⁴D²⁵EGNY²⁶IVK²⁷VNI²⁸QNG²⁹TL³⁰SAS³¹Q³²
KI³³QVT³⁴VDD³⁵PV³⁶TK³⁷PVV³⁸QI³⁹HPP⁴⁰SGAVE⁴¹YVGN⁴²MTL⁴³TCH⁴⁴VEG⁴⁵TR⁴⁶LAY⁴⁷QWL⁴⁸KNG⁴⁹RPV⁵⁰HTS⁵¹ST⁵²YS⁵³FS⁵⁴
PQ⁵⁵NN⁵⁶TL⁵⁷HI⁵⁸AP⁵⁹VT⁶⁰KED⁶¹IGN⁶²YSCL⁶³VRN⁶⁴PV⁶⁵SE⁶⁶MES⁶⁷DI⁶⁸IMP⁶⁹II⁷⁰YGP⁷¹YGL⁷²QV⁷³NSD⁷⁴KGL⁷⁵KV⁷⁶GE⁷⁷VFT⁷⁸V⁷⁹
DL⁸⁰GE⁸¹AIL⁸²FD⁸³CSAD⁸⁴SH⁸⁵PPNT⁸⁶YS⁸⁷WIR⁸⁸RTD⁸⁹NT⁹⁰TY⁹¹II⁹²KH⁹³GP⁹⁴RL⁹⁵EV⁹⁶ASE⁹⁷KVA⁹⁸QKT⁹⁹MD¹⁰⁰YV¹⁰¹CC¹⁰²AY¹⁰³NN¹⁰⁴IT¹⁰⁵
GR¹⁰⁶QDE¹⁰⁷TH¹⁰⁸FT¹⁰⁹VI¹¹⁰ITS¹¹¹VGLE¹¹²KLA¹¹³QK¹¹⁴GS¹¹⁵LS¹¹⁶PL¹¹⁷AS¹¹⁸IT¹¹⁹GIS¹²⁰FL¹²¹II¹²²SM¹²³CL¹²⁴LL¹²⁵FL¹²⁶WK¹²⁷KY¹²⁸Q¹²⁹PY¹³⁰K¹³¹VI¹³²K¹³³Q¹³⁴
LE¹³⁵GR¹³⁶PET¹³⁷EY¹³⁸RKA¹³⁹QTF¹⁴⁰SG¹⁴¹HED¹⁴²AL¹⁴³DD¹⁴⁴FG¹⁴⁵IY¹⁴⁶EF¹⁴⁷VA¹⁴⁸FP¹⁴⁹DV¹⁵⁰SG¹⁵¹VS¹⁵²RIP¹⁵³SR¹⁵⁴SV¹⁵⁵PAS¹⁵⁶DC¹⁵⁷VSG¹⁵⁸QD¹⁵⁹LD¹⁶⁰HST¹⁶¹V¹⁶²
YE¹⁶³VI¹⁶⁴QH¹⁶⁵IPA¹⁶⁶QQ¹⁶⁷DH¹⁶⁸PE¹⁶⁹

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,
276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,
239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 113

GCAAGCGGCGAAATGCGCCCTCCGGGAGTCTTGCAAGTCTCCCTGGCAGTCTGGTGCTGTT
 GCTTTGGGGTGCCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGA
 ACTGGAGAGAAGTCTGCGGAAGGAGACTGGATGATAGAAATTTTATGCCCGTGGTGCCCTGCT
 TGTCAAAATCTTCAACC CGGAATGGGAAAGTTTTTGCTGAATGGGAGAAAGATCTTGAGGTTAA
 TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC
 TTCCTACTATTTTATCATTTGTAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
 AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAAGATATTGAGCCCGTTTCATCATG
 GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA
 TCAGGACTGCGCATAACTCTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
 GTTTTGTCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
 AGATTGCCCTTTGCTCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCCTTCAAAAAAT
 TATTATCAGAATCTGCACAACCTTTGAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
 GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAGGAACAACAAGAGACTTTCACAGAATGC
 CATAAGACAACGCTCTCTGGGTCCATCATTTGGCCACAGATAAACTCTAGTTAAATTTTATAG
 TTATCTTAATATTATGATTTTGATAAAAAACAGAAGATTGATCATTTTGTGTTGGTTTGAAGTG
 AACTGTGACTTTTTTGAATATTGACGGGTTCACTCTAGATTGTGATTAATTTGAAGAGTCTA
 CATTCAAGACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG
 GTTTAAATAGTTCTCTAATTTTGAAAAACTGTCGCAAGCAATAAGATTATGTATATTTGT
 TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAAATTACATTTCCCAAGTATTGCAATTAT
 TGAGGTATTTAAGAAGATTATTTAGAGAAAAATATTTCTCATTGATATAATTTTCTCTG
 TTTCACTGTGTGAAAAAAGAAGATATTTCCATAAATGGGAAGTTTGCCCATTTGTCTCAAG
 AAATGTGTATTTCACTGACAATTTCTGGTCTTTTTAGAGGTATATTTCAAAAAATTTCCCTGT
 ATTTTATAGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA
 TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTAAGTTCATGGTATTCTCTTGATTC
 CAACAAGGTTTGATTTTCTCTGTATTTTCTTACTTACTATGGGTACATTTTATTTTTT
 CAAATTTGGATGATAATTTCTTGGAACATTTTTTATGTTTTAGTAAACAGTATTTTTTGT
 GTTTCAACTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAATTT
 TTGGCCACTTTTTAGATTTTACATCATTTCTTGCTGAACCTCAACTTGAAATTTGTTTTTTT
 TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA
 TTTTACATTTTGAAAAATCAAGAAGCTTAATATAAAAGTTTGCAATTTCTACAGGAAAAAG
 CATCTTCTTGATATATGCTTAAATGTATTTTTGTCTCATATACAGAAAGTTCTTAATTGAT
 TTTACAGTCTGAATGCTTGATGTTTTAAATATAAACATTTTTATATTTTTTAAAGACAA
 ACTTCATATTATCCTGTGTTCTTTCTGACTGGTAATATTGTTGGGATTTACAGGTAAAA
 GTCAGTAGGATGGAACATTTTGTGTTATTTTACTCCTTAAAGAGCTAGAATACATAGTTTTT
 CACCTTAAAGAAGGGGAAAAATCATAAATACAATGAATCACTGACCATTCAGTAGTAGAC
 AATTTCTGTAATGTCCTCTTTCTTCTAGGCTCTGTTGCTGTGTAATCCATTAGATTTACAG
 TATCTGAATATACAAGTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCATT
 AAAGAGTTTGGATGTGAACCTTGATGCTATGAAAAAATCTCTAAGCACAAAAATAAACCT
 TTCTAACCACTTCATTAAGCTGAAAAAATAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 114

MAPSGSLAVPLAVLVLLWGWTHGRRSNVRVITDENWRELLEGDWMI EFYAPWCPACQNL
QPEWESFAEWGEDLEVNI AKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL
ATLFSGLLLGLCMIFVADCLCPSKRRRPQFPYPYPSKLLSESAQPLKKVEEQEADEEDVSE
EEAESKEGTNKDFPQNAIRQSLGFSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTT
 CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCAGTGCGGACAGCGGATTGGAAGAGCGGG
 AAGTCTCTGGCCGAGAGCAGTGTGACACTTCCTCTGTGACCATGAACTCTGGGTGTCTGC
 ATTGCTGATGGCCTGGTTTGGTGTCTGTGAGCTGTGTGCAGGCCGAATTCTTACCTCTATTG
 GGCACATGACTGACCTGATTATATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAAGAGTACATC
 CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC
 TAGCAAGTCAGCTGCTGATGTCTGAGGGCTACCTGGCTCACCTGTGAAATGCCTACAACTGG
 TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTCTGCAGGACTCAGCTGCA
 GGTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCACTGATGAGGACGAGATAGG
 AGCTGCCAAAGCCTGTATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAAATTCCA
 GAGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG
 GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATACGGTGTGTGGATGGAGCAGGTGCT
 AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA
 GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCTGCTC
 TCCCTTGACCCAAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA
 GGAAGAGAGAGAAAAACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAAAGGCA
 TCTATGAGAGGCTGTGGACTACCTGCCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGTGG
 GAGGGTGTCAAAGCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA
 CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCGCACA
 TCGTCAGGTACTACGATGTATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA
 CCTAAACTTGCACGAGCCACCGTTCTGTGATCCCAAGACAGGAGTCTCTACTGTCTGCCAGCTA
 CCGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCTGTGTGGCCCGAGTAAATC
 GTCGGATGCAGCATATCAAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAAT
 TATGGAGTGGGAGGACAGTATGAACCCGCACTTCGACTTCTCTAGCGCAGCTTTTGACAGCGG
 CCTCAAAACAGAGGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG
 GTGGTGCCACCGTCTTCCCTGATCTGGGGCTGCAATTTGGCCCTAAGAGGGGTACAGCTGTG
 TTCTGGTACAACCTCTTGGGAGCGGGGAAGGTGACTACCGAACAAAGACATGTCTGCCCTGCC
 TGTGCTTGTGGGCTGCAAGTGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
 TGAGACCTTGTGGATCAACAGAAGTTGACTGACATCCTTTTCTGTCTTCCCTTCCCTGGTCT
 CTTACAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCTTTGTATGTTCCATACAGGCT
 GATTTTGTGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT
 GTGACTGAAGTCCCAGCCCTTCATTAGCCTGTGCCATCCTTGGCCCCAAGGCTAGGATCA
 AAGTGGCTGCAGCAGATTAGCTGTCTAGCGCTTAGCAAGGTGCTTTGTACCTCAGGTGTT
 TTAGGTGTGAGATGTTTTCAGTGAACCAAAGTCTGTATACCTTGTTTACATGTTTGTTTTAT
 GGCATTTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAAAA

100
 90
 80
 70
 60
 50
 40
 30
 20
 10
 0

FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDLVLQDSAAGFIANLSVQRQFFP
TDEDEIGAALKMRLQDTRYRLDPGTISRGEPLGTTYQAMLSVDDCFGMGRSAYNEGYYHTV
LWMEQVLKQLDAGEEATTTKSQLDYLSYAVFQLGDLHRALELTRRLSLDP SHERAGGNLR
YFEQLLEEREKTLTNQTEAELATPEGIYERPVLYLPERDVYESLCRGEVGLTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYVDVMSDEEIERIKEIAKPKLARATVRDPKTF
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS
RRPFDGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR
TRHAACPVLVGCKWVSNKWFHERGQEFLRPGSTEVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 117

GCAGTATTGAGTTTACTTCCTCCTCTTTTGTAGTGAAGACAGACCATAATCCCGACTGTGAGTGAAATTGATTGT
 TTCATTATTATACCGTTTGGCTGGGGGTAGTTCGACACCTTCACAGTTGAAGAGCAGGCAGAGGAGTTGTGA
 AGACAGGACAATCTCTTGGGGATGCTTGGGGATGCTGGTCTCGAAGCCAGCGGGCTTGCTGTGCTTTGGCCTCATGACCC
 CAGGTTCTCTGGTTAAAACTGAAAGCCTACTACTGGCCTGGTGGCCATCAATCCATTGATCCTTGAGGCTGTGCC
 CCTGGGGCACCCACTGGCAGGGCTTACCACCTGCGACTGAGCTCCCTGTGGCTCTGCTGCGGCCAGCGCTTC
 CCTCATCTTAGGGCTGTCTCTGGGGTGCAGCCTGAGCCTCCTCGGGGTTTCTTGATCCAGGGGGAGGGAGAAG
 ATCCCTGTGTCTGAGGCTGTAGGGGAGCGAGGAGGCCACAGAAATCCAGATTCCAGAGCTCGGCTAGACCAAAGTG
 ATGAAGACTTCAAAACCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC
 GGTACATCCAGACAGAGCTGGGCTCCCGTGAAGCGTTGCTGGTGGCTGTCTGAGCCTCCCGAGCTACATGTCCA
 CTTTGGCCTGAGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTACTCTACTTCACTGGGCAGCGGGGG
 CCGGCTCCAGCAGGATGCAAGTGGTGTCTCATGGGGATGAGCGGCCGCTGGCTCATGTGAGAGACCCCTGC
 GCCACCTTCAACACACACTTTGGGGCCGACTACGACTGGTTCTTTCATGTCAGGATGACACATATGTGCAGGCC
 CCGCCTGGCAGCCCTTGTGCGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCGGGCAGAGGAGTTTCATTG
 GCGCAGGCGCAGGCCCCGTACTGTCTAGGGGGCTTTGGCTACCTGTGTGTCAGGAGTCTCCTGTCTGTCTGC
 GGGCACAATCTGGATGCTGCTGCCAGGAGACATTTCTAGTGCCCTCTGACGAGTGGCTTGGACGCTGCCCTATTG
 ACTCTCTGGGGCTCGGCTGTGTCTCAGACCAAGGGCCAGTATCGCTCATTTGAACTGGCCAAAATAGGG
 ACCCTGAGAAGGAAGGAGCTCGGCTTCTCTGAGTGCTTCCGCGGTGCAACCTGTCTCCGAAGGTACCCCTCATGT
 ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGCTTACAGTGAATAGAACAATGCAAGGCTCAGA
 TCCGGAACCTGACCGTGTGACCCCGAAGGGGAGGCGAGGCTGAGCTGGCCGTTGGGCTCTCTGCTCTTTCA
 CACCACATCTTCGCTTTGAGGTGTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCTGTGCAGATGGG
 CTCCCAAGTGCCCACTACAGGGGGCTAGCAGGGCGAGCGTGGTGATCGTTGGAGACTGCCCTGGAGCAGCTCA
 ATCGGCTATCAGCCCGCTTGGCTTCCAGAAGCAGCGACTGTCTCAACGGTCTTCGGCCTTTCGACCCAGCAC
 GGGGCTAGGAGTACACCTGGACCTGTCTTGGAAATGTGTGACACAGCGTGGGCACCGCGGGGCTTGGCTCGCA
 GGGTCAGCTGCTCGGCCACTGAGCGGGTGGAAATCCTACCTATGCCCTATGTCACTGAGGGCCACCCGAGTGC
 AGCTGTGCTGCCACTCTCGTGGCTGAAGCTGTGACGCCCCGCTTCTCTGAGGCGTTTGCAGCCAAATGTCT
 TGGAGCCACAGAAATGCAATTGCTCACCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCGCTGGAGCTCCAG
 ACCATTCTTTGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGCTACCTGGGACGAGGCTGGGCTGGCTCG
 CTGTGCGAGCAGAGGGCCCTTCCAGGTGCGACTCATGGAGCTGCTCTGAGAAGAAGCACCTGTGGACACTCTCT
 TCTTCTTACCAACGCTGTGGACAAGGCTGGGCCGGAAGTCTCAACCGCTGTGCGATGAATGCCATCTCTGGCT
 GGCAGGCTTCTTTCCAGTCAATTTCCAGGAGTTCAATCCTGCCCTGTCAACACAGAGATCACCCCCAGGGCCCC
 CGGGGCTGGCCCTGACCCCCCTCCCTCCTGTGTGCTGACCCCTCCGGGGGGCTCTATAGGGGGGAGATTG
 ACCGGACGGCTTCTGCGAGGGGCTGCTTTCACAAACGCTGACTACTGGCGGCCGAGCCCGCTGGCAGGTGAAC
 TGGCAGGCGCAGGAAGGAGGGAAGCCCTGGAGGGGCTGGAGGTGATGATGTTTTCTCCGGTTCTCAGGGCTCC
 ACTCTTTCCGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAAGCCACCGCTCAGTGAAG
 AACTCTTACACCGCTGCGCCCTCAGCAACCTGGAGGGGCTAGGGGGCGTGCCAGCTGGCTATGGCTCTCTTG
 AGCAGGACAGGCCCCAATAGCACTTAGCCCCCTGGGGGCCCTAACTCATACCTTCTCTTTGCTCTGCTCAGCC
 CCAGGAGGGCAAGCAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTAAATATGAAATGTTATTAA
 ACATGTCTTCTGCC

FIGURE 118

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQEGEGDPCVEAVGERGGPQNPDSSRARLD
 QSDDEFKPRIVPYRDPNPKYKKVLRTYIQTGLGSRERLLVAVLTSTRATLSTLAVAVNRTV
 AHHPRLLYFTQGQARAPAGMQVVS HGDERP AWMSETLRHLHTHFGADYDWFFIMQDDTY
 VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRLLLLRLRPHLDGCRG
 DILSARPDWELGRCLIDSLGVCVVSQHGGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
 GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPGEAGLSWPVGLPAPFTPHSRFEV
 LGWDYFTEQHTFSCADGAPKCP LQGASRADVGDALETALEQLNRRYQPRLRFQKQRLNLNGYR
 RFDPARGMETLDDLLEECVTQRGHRRALARRVSLRLPLSRVEILPMPYVTEATRVQVLPL
 VAEAAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDFLGVKAAAAEELERRY
 PTRLAWLAVRAEAPSQVRLMDVVSKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFPP
 VHFQEFNPALSPQRSPPGPPGAGDPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLA
 RARLAGELAGQEEEEALEGLEVMDFLRFSGHLHFRAVEPGLVQKFSRLRDCSPRLSEELYHR
 CRLSNLEGLGGRAQLAMALFEQEANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,
 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
 558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

121
125
342
346
319
323
464
468
64
68
150
154
322
326
331
337
368
372
385
389
399
403
409
413
473
477
729
733
748
752
736
743
19
25
23
29
136
142
397
403
441
447
544
550
558
564
651
657
657
663
672
678
14
25
247
250

FIGURE 119

CGGAGTGGTGCGCCAACTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCTGTCCCCAAGCC
 GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
 AAGCATTCTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAATAGAA
 TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAAATTCA
 GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTGAGTATACTGTATTATCCTTGTAACC
 CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACTGTGACAAAGCAG
 AGTTCTTCAGTTCTGAAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG
 TGGTTAATGATGAGAAAAAGCTTACAAATACGCCCTTTGATAAGTATAGAGACCAATACAACATG
 GTTCTTCTTGACGCCCCACTACGTTTGCTATCATTTGAAAACCTAAAGTATTTTTTGTTAA
 AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
 GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
 CAATATCCAGAAAAGTGTCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC
 AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA
 AAAGATGTATTTAATACCAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC
 CAACCAGGTAGTAGAAGGCTGTGTTCAGATATGGCTGTTACTTTTAATGGACTGACTCCAA
 ATCAGATGCATGTGATGATGATATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTCAAT
 GATGCATTGGTTTTCTTACCTCCAAATGTTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG
 TGAATATGATCTTTGTATAGGACGTGTGTGTGTCATTATTTGTAGTAGTAACTACATATCCAA
 TACAGCTGTATGTTTTCTTTTCTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAG
 TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTCTTTAAACACATGAACATTGTAAATG
 TGTGGAAGAAGTGTTTTAAAGATAATAATTTGCAAATAAACTATTAATAAATATATAT
 GTGATAAATTCTAAATATGAACATTAGAAATCTGTGGGCACATATTTTGTGATGGTT
 AAAAAATTTTAAAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
 TGTGATTAAAGTAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
 AAGCCTCCCCAAGTTCCAATGGATTGCGCTTCTCAAATGTACAATAAGCACTAAAGAAA
 ATTAAGTGAAAGTTGAAAAAT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME
LSKSFVRVYCIILVKPKDVSLSAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFAI IENLKYFLKKDPSQPFYLGHTIKSGDLEYVGMEGG
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT
KSVGLS I KEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL
PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

[illegible][illegible]

FIGURE 122

MNSSKSSETQCTERGCFSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALS LKNC SAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLED CATMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

FIGURE 123

GGGACTACAAGCGCGCCGCGCTGCCGCTGCCCTCAGCAACCTCGACATGCGCGCTGAGGCGGCCACCGCGAC
 TCCGCTCTGCGCTCGGCTGCTGACTTCTTCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC
 TCAATCCAGCAATCGCAACCCAGTGGTACAGGAATTTGAAAGTGTGGAACCTGTCTGCATCATTACGGATTCCG
 AGACAAGTGACCCAGGATCGAGTGGAGAAAATTAAGATGAACAAACCAATGTGTGTTTTTGACACGAA
 TTTCAAGGAGACTTGGCGGGTCTGTGCAGAAAATCTGGGGGAGACATCCCTGAAGATCTGGAATGTGACACGGAG
 ACTCGAGCCCTTTATCGCTGTGAGTGTGCTGCTGAAATGACCGCAAGGAAATGTATGAGATGTGTGATCGAGTTAA
 CTGTGCAAGTGAAGCGCAGTGACCCCTGTCTGTAGAGTGC CGAAGGCTGTACCAAGTGGCAAGATGGCAACACTGC
 ACTGCCAGGAGATGAGGGCCACCCCGGCTCACTACAGCTGGTATCGCAATGATGTACCATGCCACGGATT
 CAGAGCCCAATCCGAGATTTCCCAATTTCTTTCCACTTAACTCTGAAACAGGCCTTTGGTGTTCACGTGTG
 TTTCAAGAGCAGCTCTGGGCGACTACTGCTGCTTCAATGACGCGGCTCAGGCTCAGCCAGGTGTGAGGAGCAGG
 AGATGGAAGCTATGACTCTGAACATGGCGGAATTTATGGGGGGTCTCGGTGTCTCTGCTGTGCTGCTGCTGA
 TCACGTGTGGGCATCTGCTGTGCATACAGAGCTGGCTACTTCACTCAACAATAAACAGGATGGAGAAAGTTACAAGA
 ACCCAGGGAACAGATGGATTTAACTACATCCGCACTGACGAGGAGGGCGACTTCACACAAAGTCATCGTTTTG
 TGATCTGAGACCCCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAACTCTCTGCA
 GGCACGAGAGCTGATGCACTCGGACAGAGCTAGACACTCAITCAGAAGCTTTTCTGTTTGGCCAAAGTTGACCA
 CTACTCTTCTTACTCTAAACAAGCCACATGAATAGAAGAAATTTTCTCAAGATGGACCCGTAATATAACCAAA
 GGAAGCGAAATCTGGGTGCTCTCACTGATGTGGGTTCCTAATCTGTTTCTGGCTGATTTCCCGCATGATATTAGG
 GTGATCTTAAAGAGTTTGTCTTACATAAACCGCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCACCACTGGTCTG
 CAGCAGCCACACAGCACCATTGTGAGATGGCGAGGTGGCTGGAAGCAGCAGCAGCGCATCCCGGCGGAAACCA
 GAAAAGGCTCTTACACAGCAGCTTACTTTCATCGGCCACAGACACCACCGCATTTCTCTTAAAGGCTCTGCT
 TGATCGGTGTTCAGTGTCCATTGTGGAGAACTTTTGTGATCAGCATTTTGTAAACCAACCAAAATCAGGAAG
 GTAAATTTGGTGTGGAAGAGGGATCTTGCTGAGGAACCTCTGTTGTCACAGGGGTGTGAGGATTTAAGGAAA
 ACCTCTGCTTATAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTGTGTTTATTTATAAAATTT
 TACATCTAAATTTTGTCTAAGGATGTATTTTGATTATTGAAAAGAAATTTCTATTTAAACTGTAAATATATTGT
 CATACAATGTTAAATAACCTATTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT
 TGGAAAATCAATAATTAAAGATATTTTACCAGGAATCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT
 CACACAAGTTTATGCTTTTTCACAAAGGAACCTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
 TAAAAATTCAGTTAAGCAATTTGAAATCAGTTTGCACTCTTCAAAAGAACTCTCAGGTGTGCTTTGAACT
 GCCTCTTCTGAGATGACTAGGACAGTCTGTACCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG
 CCACTCAGCTCTCTGGGTTGCGCCAGGCGCCCGCTCTAGCTCACTGTGCTCTGCTGTCTGCCAGGAGGCCCT
 GCCATCTCTGGGCCCTGGCAGTGGCTGTGTCCAGTGAGCTTTACTCAGCTGGCCCTGCTTCACTCAGCACAGC
 TCTCAGGTGGGCCTGCGAGGACACTGGTGTCTTCCATGTAGCGTCCAGCTTTGGGCTCCTGTAACAGACCTCT
 TTTTGGTTATGAGTGCTCTCAAAAATAGGGCCCCCAATGCTATTTTTCTTTTAAAGTTTGTTTAATTTATTTGTT
 AAGATTCTTAAGCGCAAGGGCAATTCGCAAAATCAAGTCTGTCAAGTACAATAACATTTTAAAAAGAAATGGAT
 CCCACTGTTCCTCTTCCACACAGGAAGCACCAGAGCGCCAGGCTCTGTGCGATTTCAAACCAAAACCATGAT
 GGAGTGGCGCCAGCTCCAGCTTTTAAAGAACCTCAGGTGGAGCAGCCAGGTGAAAGGCTGGCGGGGAGGAAAG
 TGAACCGCTGAATCAAAGCAGTTTCTTAATTTTGACTTTAAATTTTTCATCCGCGGAGACACTGCTCCCAAT
 TTTGGGGGCACTTGAACAACACTCAGAGGCTGTGTTCTTCAAGAGCAGGTGCTCAGCCCTCAACATGCTTTT
 GCGCTCTGACTCAGACTGAAGTGTGTAAAGCAAGGAGCTGTGTAGAAGGAGCACTTCACTGTGTGCTGGGA
 GAATGGCTCTCACTGACTACCTTTGCTTTTCACTTCCAGTCTCAGTGTCTTGGGTTTTTATATCTGTGACAGCTTTT
 AATTGTCATCACTAGCACTGTGTTGACTTTTTTAGTTGTGTGAAACACTTTGGCCGAGGCGCTGGCAGAGGCA
 GGAATGTCTCAGCAGTGGCTCAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 CCTCATCATTTGCCACCTTGTGTAGAGAGGGATGCTGCCACCTCAGCGTGGGGATTCAGCTCTGCAAGTTCT
 TCTTGTGTGTATGATGATGGTGGCTTATTTGCCCCCTCTTATACCTCAAAACCTTCTACACTAGTGCCCA
 TGGGAACAGGCTGTGAAAAGTAGAGAGAAAGTAAAGTAGAGTCTGGAGAGTACCTGCTATACTGAGACTAGA
 CGGAAAGGAATCTCGTGTATTTAAGATGTAATGTGACTCAAGACTCAGCGCATCGAGCGCATCGAGGCTGTGATTTCT
 GCCTTTGATGGATGTGCTGTGCTACAGATGCTACAGCTTTGATTAACACACCGTAATTTGGCATTTGTTTAACT
 CTCATTTATAAAGCTTCAAAAAACCA

FIGURE 124

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624
><subunit 1 of 1, 310 aa, 1 stop
><MW: 35020, pI: 7.90, NX(S/T): 3
MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTFVVQEFESVELSCIITDSQTS
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLSKIWNVTRRDSALYRCEVVARNDRK
EIDEIVIELTVQVKPVPVCRVPKAVPVGKMATLHCQSESGHPRPHYSWYRNDVPLPTDSRA
NPRFRNSSFHLNSETGLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGG
VLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI
```

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267